Regulation of lipid metabolism by GATA6:
an integrated 'omics' approach

Patrícia Raquel Dos Santos Rodrigues

Institute of Infection and Immunity

School of Medicine

Cardiff University

A thesis submitted to Cardiff University in accordance with

the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2018

Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed……(candidate)
Date……

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed……(candidate)
Date……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed……(candidate)
Date……

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.

Signed……(candidate)
Date……

"It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of light, it was the season of darkness, it was the spring of hope, it was the winter of despair."

– Charles Dickens, A Tale of Two Cities

Acknowledgments

First, and foremost, I would like to thank my supervisor Prof. Valerie O'Donnell. The work developed during this project wouldn't have been the same without Prof. O'Donnell's constant presence, the door to her office was always open whenever I had a question about my research or writing. She consistently allowed this project to be my own work but steered me in the right direction whenever she thought I needed it. I am also profoundly grateful to Prof. Phil Taylor for his support and ongoing contribution to the work developed during my PhD studies.

My research would have been impossible without the aid and support of Dr. Marcela Rosas. Her initial work helped construct this project and her support was critical for its development.

I would also like to acknowledge the members of the O'Donnell and Taylor groups, both past and present, which have contributed for this work and made it possible. Dina Fathalla, Natacha Ipseiz and Magdalena Czubala gave me invaluable help with the lentiviral work conducted herein. I would also like to thank everyone who helped me with in any way, Victoria Tyrrel, Keith Allen'Redpath, Marya Misheva, Anastasia Gketsopoulou, Dan White, Jorge Alvarez Jarreta, Maceler Aldrovandi, Christine Hinz and Sven Meckelmann.

I would like to thank the European Research Council for funding my PhD studies, as well as the British Heart foundation and the Biochemical Society for supporting laboratory and conference visits.

A very special thank you to James Burston for being an unmatched fount of knowledge, for always sharing his expertise in order to make me a better scientist, and for always being able to listen and try to help me whenever I had questions. I can't seem to find the right words to thank you enough.

I can't help mentioning my appreciation towards for my friends, for helping make this an enjoyable experience and to everyone who has ever believed in me and supported me in this journey. In particular, I would like to thank Silvia Gimeno Brias for all our adventures around Wales.

Por fim, não podia deixar de agradecer ao César Sobreira por me ter apoiado e ter feito sempre os possíveis para que eu conseguisse atingir os meus objectivos.

Summary

The transcription factor GATA6 was recently recognised as a master regulator of the phenotype and function of peritoneal resident macrophages (pResMϕ), whose deficiency results in dysregulated proliferative renewal during homeostasis and altered inflammatory responses, associated with delays in resolution.

Herein, I show using microarray analysis that mice with a myeloid deficiency of *Gata6* (*Gata6*-KO^{mye} mice) have significant changes in genes associated with lipid metabolism, in particular sphingolipids (SL). Three of the most notable alterations were the downregulation of genes responsible for the degradation of glucosylceramides (GlcCer) and sphingomyelins (SM) (*Gba2* and *Smpd1*), as well as a gene involved in the regulation of lysosomal pH, and consequently, lysosomal function. To characterise the metabolic defect, I compared the lipidome of *Gata6*-KO^{mye} and wild-type (WT) pResMϕ using a high-resolution LC-MS based global lipidomics approach and replicated my findings using targeted LC-MS/MS. SL showed a high percentage of changes (25% of all lipids levels were altered significantly) with a marked increase in *Gata6*-KO^{mye} pResMϕ. SL are important constituents of the plasma membrane in eukaryotes and, as second messengers, modulate apoptosis, cell proliferation and differentiation. Targeted lipidomics by LC-MS/MS showed that the most significantly-increased molecular species were long chain GlcCer and SM. This linked with downregulation of *Gba2*, *Ctse* and *Smpd1*, observed during the microarray analysis, and was confirmed by supressing these pathways, in RAW 264.7 cells, using knockdown (shRNA) approaches. Accumulation of SL metabolites in tissues is implicated in a plethora of patho-physiological complications such as development of neurological dysfunctions, atherosclerosis, diabetes, and heart failure, whilst in cells it correlates with defects in cell migration, cholesterol traffic and efflux, lipid transport, cell activation, proliferation and apoptosis.

Overall, I concluded that GATA6 regulates lipid metabolism in pResMϕ, on both a transcriptional and metabolic level, with particular focus on SL levels. Furthermore, there is indication that this regulation may at least in part control phenotype and function in pResMϕ through altering SL signalling. Thus, my studies propose GATA6 as a new lipid-regulating transcription factor and highlight the importance of lipid regulation in pResMϕ biology.

Table of Contents

List of figures

List of tables

List of abbreviations

Abreviation	Name
2-AG	2 arachidonoylglycerol
3KSa	3-ketosphinganine
AA	Arachidonic acid
AEA	arachidonoylethanolamide
bp	Base pairs
C1P	Ceramide-1-Phosphate
CAPP	Ceramide activated Ser-Thr phosphatase
CDase	Ceramidase
Cer	Ceramide
CerP	Ceramide 1-phosphate
CerS/LASS	(Dihydro)-ceramide synthase
CK	Ceramide kinase
CNS	Central nervous system
COX	Cyclooxygenase
Ctse	Gene encoding for Cathepsin E
CTSE	Cathepsin E
DAGL	diacylglycerol lipase
DHA	docosahexaenoic acid
DNA	Deoxyribonucleic acid
DPA	docosapentaenoic acid
ds-shRNA	Double stranded short hairpin RNA
eCB	endocanabinoid
eGFP	Enhanced green fluorescence protein
$eGFP^+$	Green fluorescence protein positive cells
EMP	Erythromyeloid progenitor
EPA	eicosapentaenoic acid
ER	Endoplasmic reticulum
ESI	electrospray ionization
FAAH	fatty acid amide hydrolase
FC	Fold-change
FT	Fourier transform
GATA	GATA-binding protein
Gata	Gene encoding for GATA binding protein

Gba2	Gene encoding for Glucosylceramidase Beta 2
GBA2	Glucosylceramidase Beta 2
GC	Gas chromatography
GCS	Glucosylceramide synthase
GFP	Green fluorescence protein
GlcCer	Glucosylceramide
GPR	G protein-coupled receptors
GSL	Glycosphingolipid
HETE	hydroxyeicosatetraenoic acid
HSC	Hematopoietic stem cells
IGF	Insulin-like growth factor
IL-1	interleukin-1
ImmGen	Immunological Genome Project
InfMϕ	inflammatory macrophages
KO	Knockout
L-type	long GATA6 isoform
LacCer	Lactosylceramide
LAMP-1	Lysosomal-associated membrane protein
LC	Langerhans cells
LIPID MAPS	Lipid Metabolites and Pathways Strategy
LOX	Lipoxygenase
LPA	lysophosphatidic acid
LPC	lysophosphatidylcholine
LPI	lysophosphatidylinositol
LPSer	lysophosphatidylserine
LT	leukotriene
LX	Lipoxin
LXR	Liver X receptor
Lyso-PLD	lyso-phospholipase D
MAGL	monoacylglycerol lipase
MaR	Maresin
MDP	Macrophage and/or dendritic cell progenitors
Mϕ	Macrophage
MOP	Monocyte precursor cells
MP	Myeloid precursors
MPS	mononuclear phagocyte system
MRM	Multiple-reaction monitoring

MS	Mass spectometry
NAPE-PLD	N-arachidonoylphosphatidylethanolamide-specific phospholipase D
NFATC1	Nuclear factor of activated T-cells
NLS	Neutral-loss scan
NPC	Niemann-Pick disease
Oct4	octamer-binding transcription factor 4
oxLDL	oxidised low-density lipoprotein
PA	Phosphatidic acid
Pal-CoASH	palmitoyl coenzyme A
PC	Phosphatidylcholine
PCR	Polymerase chain reaction
PD	Protectin
PDGF	platelet-derived growth factor
PDX	Protectin DX
PE	Phosphatidylethanolamine
PG	Prostaglandin
PI	Phosphatidylinositol
PIS	Precursor-ion scan
PKC	protein kinase C
PKH	PKB homologue
PLA2	Phospholipase A2
PPARγ	proliferator-activated receptor-γ
pResMϕ	Peritoneal tissue resident macrophage
PU1	Spi-1 Proto-Oncogene
PUFA	polyunsaturated fatty acids
qPCR	Real-time polymerase chain reaction
RA	Retinoic acid
ResMϕ	Tissue resident macrophage
RNA	Ribonucleic acid
RNAi	RNA interference
RT-PCR	Reverse transcription polymerase chain reaction
Rv	resolvin
S-type	Short GATA6 isoform
S1P	Sphingosine-1-Phosphate
Sa	Sphinganine
SFPI1	Spi-1 proto-oncogene
shRNA	Short hairping RNA

siRNA	Small interfering RNA
SK	Sphingosine kinase
SL	Sphingolipid
SM	Sphingomyelin
SMase	Sphingomyelinase
Smpd1	Gene encoding for Acid sphingomyelinase
SMPD1	Acid sphingomyelinase
SMS	Sphingomyelin synthase
So	Sphingosine
SPPase	Sphingosine phosphate phosphatase
SPT	Serine palmitoyl transferase
SRM	selected reaction monitoring
TGFβ	transforming growth factor-β
TOF	time-of-flight
TX	thromboxane
VEGF	vascular endothelial growth factor
WT	Wildtype
YPK	yeast protein kinase

Chapter 1 General Introduction

1.1 Macrophages

The theory and proof of natural, innate immune host defence, by cellular mechanisms such as phagocytosis and recruitment of phagocytes during acute and chronic inflammation, was first described by Ilya Mechnikov, a Russian comparative zoologist known as the father of innate immunity and discoverer of the significance of phagocytosis in development, homeostasis and disease[1]. In 1883, he observed specialized phagocytic cells, microphages and larger macrophages (Mϕ), consuming pathogens during inflammation, which led him to develop the theory of immunity by phagocytosis[1,2]. Since then, the interest in Mϕ biology has steadily increased, and it has become evident that the multiplicity of Mϕ functions play a vital role in tissue development, homeostasis and defence[3,4].

Phagocytic cells, can be divided into two groups: the polymorphonuclear phagocytes (granulocytes) and the mononuclear phagocytes. Mononuclear phagocytes comprise tissue circulating monocytes, resident Mϕ (ResMϕ), promonocytes, and their precursor cells in the bone marrow[5]. More recently, the concept that has emerged that Mϕ represent an evolutionarily ancient, disseminated, homeostatic system, on a par with the nervous and endocrine systems[6]. They are present in all vertebrate tissues, from mid-gestation throughout life, representing a tissue-distributed mononuclear phagocyte system (MPS)[5,7]. They maintain homeostasis by responding to physiological changes and external insults, which results in a wide range of reactions, as phagocytes in defence against microbes and in clearance of dead and senescent cells, in addition to providing regulatory and repair functions[8]. During organogenesis, Mϕ orchestrate tissue remodelling and relay long-range signals between non-immune cells[9,10]. They do this by regulating pigmentation in zebrafish skin[11], and fostering spermatogonial differentiation in the murine testis before birth[12]. During puberty[13], they achieve it by promoting branching morphogenesis in the developing kidney[14], carving the developing cornea[15] and wiring neurons in the developing brain[16], among others. Mϕ also have a crucial role during erythropoiesis in the foetal liver by enucleating primitive erythrocytes[17]. Murine heart Mϕ are responsible for tissue regeneration after

birth[18]. Last, they also control the inflammatory responses required for tissue repair and wounding during adulthood[19].

The past few years have witnessed an immense advance in our understanding of the origins and distinct functions of monocyte/Mϕ subsets (Figure 1.1) as well as their differential roles in health and disease[8,20,21]. In the classical view, Mϕ develop from self-renewing hematopoietic stem cells (HSC) in the bone marrow via blood monocyte intermediates[22]. In the 1960s, Von Furth and Cohn proposed that all Mϕ, not only those in inflammatory foci but also those in tissues in the steady state, are derived from monocytes, which differentiate via promonocytes from monoblasts originating in bone marrow[5]. Hence, Mϕ and Mo were classified as part of the MPS, and ResMϕ were instead considered to be short-lived, terminally differentiated cells[23,24]. This model was broadly accepted until recent fate-mapping studies demonstrated that some adult ResMϕ populations, such as Kupffer cells, Langerhans and microglia, develop from embryonic progenitors independent of HSCs and can self-renew[25]. Local proliferation of ResMϕ can assure homeostatic maintenance and dramatically increase cell number upon challenge[25,26]. ResMϕ are of embryonic origin and are maintained throughout life autonomously from conventional haematopoiesis by self-renewal and/or longevity (Figure 1.1)[27–29]. In contrast, Mϕ infiltrating the tissues during an inflammatory response are derived from blood monocytes[30,31].

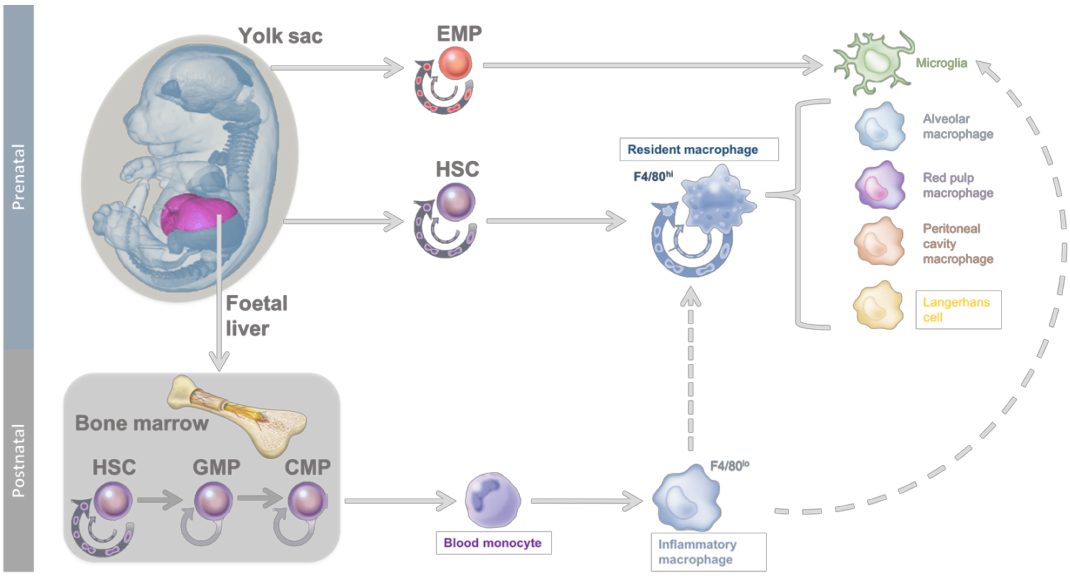


Figure 1.1 Mϕ origin and self-renewal.

The figure illustrates the precursor cells of different ResMϕ. Microglia derive from immature, erythromyeloid progenitors (EMPs): that is, stem cells that are formed during primitive haematopoiesis in the yolk sac in mice. By contrast, other ResMϕ arise from definitive haematopoiesis, first in the aorta–gonad–mesonephros region or later, in the foetal liver. Postnatally, monocytes are formed in the bone marrow. Haematopoietic stem cells (HSCs) generate monocytes from myeloid precursors (MPs) and Mϕ and/or dendritic cell progenitors (MDPs).

1.1.1 Tissue-resident Mϕ

Mϕ activities such as mediating host protection versus tissue injury, are kept in a fine balance. Mϕ phagocytosis, clearance and secretion contribute to innate and adaptive defences against infection and support inflammation, whilst the same processes with different signals restore tissue homeostasis and promote repair[32,33]. They are a versatile, heterogeneous group of cells with similar functions and phenotypes, crucial for tissue defence and homeostasis. Mϕ are adaptable and have the ability to perform trophic and defence functions, both locally and systemically[34]. Lymphocytes and Mϕ interact on different levels, initiating cellular and antibody immune responses, while executing effector roles which, if excessive or perpetuated, lead to tissue destruction[35].

Although ResMϕ are anatomically distinct from one another, and have different transcriptional profiles and functional capabilities, they are all essential for the maintenance of tissue homeostasis[36]. For example, in the lungs, ResMϕ prompt tissue-specific responses such as immune surveillance of inhaled pathogens[37]. Nevertheless, chronic insults subvert their reparative and homeostatic functions, leading to a causal association between Mϕ with disease states, such as fibrosis, obesity and cancer[36]. During malignancy, ResMϕ play an important role in promoting the survival, growth and spread of tumour cells[35,38], whereas in the CNS ResMϕ (microglia) have been shown to induce astrocyte activation[39].

In the past few years, research has highlighted that the tissue environment itself controls the ResMϕ phenotype, by influencing the expression of specific genes[40]. Thus, ResMϕ have specialized functions dictated by their physical locations and distinct gene expression profiles. This suggests that ResMϕ will have different responses to inflammation, depending on their location[41]. However, there is still a lot that is not yet fully understood concerning ResMϕ tissue diversity/plasticity, categorization of specific tissue environments, and the role of different stimuli in ResMϕ metabolism, in particular lipids. New insights

on how Mϕ subsets can be manipulated to control the inflammatory environment in a tissue-specific manner for therapeutic purposes are now required[2].

Despite the fact that we currently know of a variety of ResMϕ and they have been the focus of numerous studies, it is still unclear whether all ResMϕ are equally capable of self-renewal or if 'progenitor cells' exist within the population[42]. In the lungs, it has long been thought that following depletion of alveolar Mϕ, the majority of repopulation occurs by stochastic cellular proliferation[43]. More recently, studies in the dermis[44], gut[45], and heart[46,47] have led to a change of thinking. Clonal self-renewal among Langerhans cells (LC) suggest the existence of a multitude of LC progenitors or mechanisms that enable subsets of LCs to acquire potentially transient and limited progenitor-like properties[48]. Embryonic precursor cells were shown to seed the intestinal mucosa and demonstrated extensive *in situ* proliferation during the neonatal period. However, at the time of weaning, these cells become replaced by monocytes that later differentiate into mature, inflammatory Mϕ[45]. Furthermore, in the heart, embryo-derived Mϕ show declining self-renewal with age and are progressively substituted with monocyte-derived Mϕ, even in the absence of inflammation[46,47]. Interestingly, in these tissues, the decline of self-renewal and subsequent substitution by monocyte-derived Mϕ increases with age and accompanies the loss of proliferation within the embryonic population.

ResMϕ are thought to be long-lived cells that arise from a transient hematopoietic wave of erythro-myeloid progenitors (EMPs) emerging in the yolk sac. There are currently three possible origins for ResMϕ:

(i) EMPs emerging from the yolk sac, which are c-Myb- independent, and lead to the formation of yolk sac Mϕ[49].

(ii) c-Myb-dependent, arise when EMPs seed the foetal liver where they generate foetal monocytes that differentiate into primitive Mϕ, and seed tissues, giving rise to microglia in the brain[49,50].

(iii) HSC-derived monocytes, emerging from the foetal liver, contribute to the long-lived ResMϕ pool established at birth[3].

Nevertheless, it is still unclear, to which extent a self-renewing resident population of Mϕ is maintained through life[27].

a) ResMϕ Heterogeneity

Mϕ express a plethora of sensing molecules, as well as scavenger, hormone, cytokine and pattern recognition receptors, that allow them to oversee tissue microenvironments and act as sentries for infection and tissue damage[51]. This sensing capacity, renders them plastic and remarkably adapted to their specific local environments[52,53]. Moreover, ResMϕ can perform tissue specialised functions that reflect their phenotypic diversity, further suggesting that tissue-derived signals may control the development of tissue-specific phenotypes[54,55]. These signals are largely unknown, with few exceptions, such as, the RANK signalling pathway in osteoclasts (which share a lineage with Mϕ)[51,56].

In order to fully comprehend ResMϕ heterogeneity it is imperative to first understand their detailed transcriptional regulation[36]. It is increasingly appreciated that, the development of ResMϕ phenotypes is controlled by distinct transcriptional master regulators (Figure 1.2)[57]. In several cases (described below), the deletion of transcription factors that regulate tissue-specific transcription programs, results in ablation of Mϕ subsets. For example, *Spi-C* deficient mice exhibit a defect in the development of red pulp Mϕ, that fail to phagocytose red blood cells efficiently, and develop an iron overload localized selectively, that drives inflammation[58]. Liver X receptor (LXR) deficient mice are defective in the generation of marginal zone and metallophilic Mϕ, causing abnormal responses to blood-borne antigens[59]. Nuclear factor of activated T-cells (NFATC1) deficient embryonic stem cells fail to differentiate into osteoclasts[60],whilst loss of *Spi-1* proto-oncogene (SFPI1) (also known as PU.1) results in complete depletion of $CD11b^{+}F4/80^{+}$ Mϕ, including those derived from the yolk sac[61]. The loss of these cell populations, when specific transcription

factors are knocked-down, makes it extremely difficult to study the mechanisms by which they are regulated. It does however indicate that these transcription factors are essential for ResMϕ survival/development[2].

In 2012, the Immunological Genome Project (ImmGen) assessed ResMϕ gene expression in several mouse organs (lung, liver, peritoneum, spleen and brain). This showed a significant diversity in the transcriptional profiles based on tissue of origin, including transcripts that were universally associated with mature ResMϕ. In particular, TCEF3, C/EBPα, BACH1, and CREG-1 were among the top transcriptional regulators predicted to regulate core ResMϕ-associated genes[57]. The ImmGen algorithms predicted additional transcription factor mRNAs that were strongly associated with single Mϕ populations, such as, Spi-C for red pulp Mϕ, confirming a regulation that was already known[58], and GATA6 as a regulator of pResMϕ[62]. They also found that pResMϕ specific genes could either be GATA6-dependent or -independent. Accordingly, tissue-specific Mϕ phenotypes can be described by an arrangement of several transcription factors, where each one controls distinct transcription functional programs[51]. This work underscored the shared and distinctive characteristics of different ResMϕ and provided significant evidence to distinguish these cells from dendritic cells.

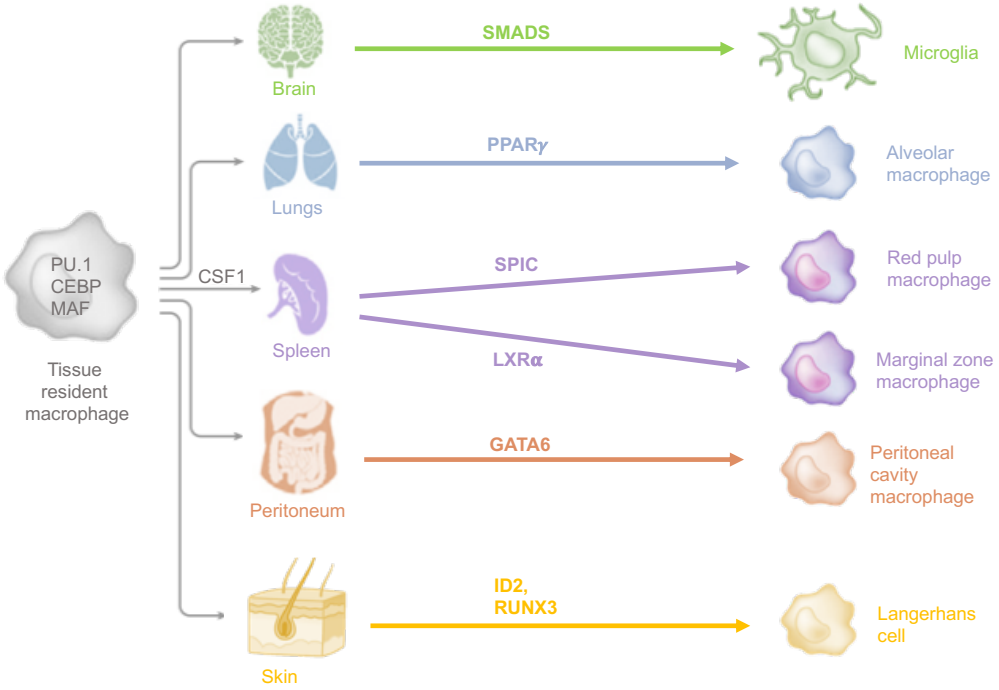


Figure 1.2 Regulation of Mϕ development and function in peripheral tissues.

The figure depicts how the expression of different transcription factors and the local enviromnment mould the phenotype and function of ResMϕ. In the brain, incoming yolk sac-derived cells are exposed to locally expressed transforming growth factor-β (TGFβ), which drives SMAD phosphorylation and the expression of genes that are unique to microglia. In the lungs, proliferator-activated receptor-γ (PPARγ), drives differentiation into alveolar Mϕ. In the spleen, SPIC expression controls differentiation and maintenance of red pulp Mϕ. In the marginal zone of the spleen, liver X receptor-α (LXRα) is responsible for Mϕ maintenance. In the peritoneum, GATA-binding protein 6 (GATA6), promotes the differentiation of resident Mϕ. In the epidermis ID2 and RUNX3 are responsible for Langerhans maintenance.

b) The peritoneal cavity: a model for tissue-specific transcriptional control

The peritoneal cavity has long been known to be home to a mix of immune cells with important functions in monitoring visceral organs and related mesothelium. pResMϕ are characterised by a high expression of F4/80 and share a transcriptional similarities with other ResMϕ[57]. However, there is a less abundant and phenotypically distinct group of phagocytes also present in the peritoneum, characterized by lower F4/80 and higher MHCII expression[63]. In 2010, the murine 7/4 antigen was used to distinguish between monocytes/monocyte-derived cells recently recruited to a site and the resident cells already present. Thus, the F4/80^{lo}MHCII^{hi} cells were characterised as potential Mo derived with an immature phenotype[64].

Recently, Bain *et al.* showed that the peritoneal cavity is initially seeded by embryo-derived Mϕ. After birth, these are displaced by bone-marrow derived monocytes which widely phenocopy the embryo derived cells, displaying extensive proliferative capacity[42]. This disproved the previous view that *in situ* self-renewal is synonymous with embryonic origin. Moreover, it was established that F4/80^{lo}MHCII^{hi} peritoneal Mϕ are intermediaries between newly arrived monocytes and mature F4/80^{hi} pResMϕ. This indicates that homoeostasis of pResMϕ occurs as a combination of self-maintenance and renewal from bone-marrow-derived monocytes (BMDM)[42].

pResMϕ are among the best-studied ResMϕ in terms of cell biology, development, and inflammatory responses[51,65,66]. They play essential roles in the control of infections and inflammatory pathologies when studied in model systems[67,68]. Nonetheless, the tissue-specific function of pResMϕ as a cell population remains unclear[51].

Since the ImmGen study was published, pResMϕ have been used as an experimental model to investigate tissue-specific functions and external cues that control their specific transcriptional program. Okabe *et al.* performed a gene expression analysis where the transcription factor GATA6 was once more

identified as a regulator of a tissue-specific gene expression program in pResMϕ[51]. Separately, a microarray analysis performed by Prof. Phil Taylor's group (Cardiff University), on monocyte and Mϕ populations, with distinct lineage origins[63], including several isolated during peritoneal inflammation, enabled the characterisation and comparison of transcriptional profiles of pResMϕ and inflammatory Mϕ (InfMϕ)[69] (see Table 1.1 for full list of cells analysed). This analysis defined the ResMϕ-restricted transcriptional profile and pResMϕ-specific probes (cluster 15). Among these probe sets was the zinc finger transcription factor GATA6, which was selectively expressed in pResMϕ, CD11b^{hi}F4/80^{hi}MHCII^{lo} cells (Figure 1.3). *Gata6* was one of the many targets linked to proliferation and enriched in pResMϕ, and its deficiency led to dysregulated proliferative renewal during homeostasis and in response to inflammation[69].

Of relevance, *Gata6* is selectively expressed in pResMϕ and alveolar Mϕ[57], and its expression is induced by tissue-derived retinoic acid (RA)[51](Figure 1.4A). Additionally, Okabe et al found that when the precursor for RA (vitamin A) is depleted from the diet, mice present a decrease in *Gata6* expression and pResMϕ levels[51] (Figure 1.4B). However, when stimulated with retinoic acid, pResMϕ from vitamin A depleted mice exhibit an increased expression of *Gata6* and pResMϕ-specific genes[70](Figure 1.4C). Thus showing that *Gata6* expression can be induced by vitamin A signalling in foetal liver derived Mϕ[51]. However, this has not yet been demonstrated in monocytes or other Mϕs from adult mice, including Mo-derived Mϕ recruited to the peritoneum during thioglycollate-elicited inflammation[20]. This indicates that the heterogeneity of ResMϕ might be due to a transcriptional program triggered by tissue-derived signals. Moreover, in the absence of GATA6, the CD11b^{hi}F4/80^{hi}MHCII^{lo} Mϕ population in the peritoneum is greatly contracted, but the CD11b^{lo}F4/80^{lo}MHCII^{hi} population that typically lacks *Gata6* expression remains at normal frequency, indicating that GATA6 is not essential for pResMϕ development but is required for their maintenance in the peritoneum[51,69,71].

In the next section I will introduce the GATA6 transcription factor in greater detail, with particular focus on the aims of this thesis.

Table 1.1 Purification criteria for monocyte and Mϕ populations (Rosas, M. *et al.* Science 344, 645–648 (2014)).

Cell type	Zymosan	Purification Criteria	Description
Bone marrow monocytes	**0 hours**	Ly-6B^{hi}, Ly-6G^{-}	Naïve bone marrow monocytes
Inflammatory macrophages	**4 hours**	Ly-6B^{hi}, Ly-6G^{-}	Early recruited monocytes
Inflammatory monocytes/ Mϕ	**18 hours**	Ly-6B^{+}, Ly-6G^{-}	Late recruited monocytes/InfMϕ
Mϕ	**3 days**	F4/80$^{+/hi}$, CD11b$^{+/hi}$	Mixed ResMϕ/InfMϕ
Ly-6B^{+}Mϕ	**7 days**	F4/80^{hi}, Ly-6B^{+}	Ly-6B^{+} InfMϕ
Ly-6B^{+}Mϕ	**7 days**	F4/80^{hi}, Ly-6B^{-}	Predominantly ResMϕ
ResMϕ	**0 hours**	F4/80^{hi}, CD11b^{hi}	Naïve ResMϕ

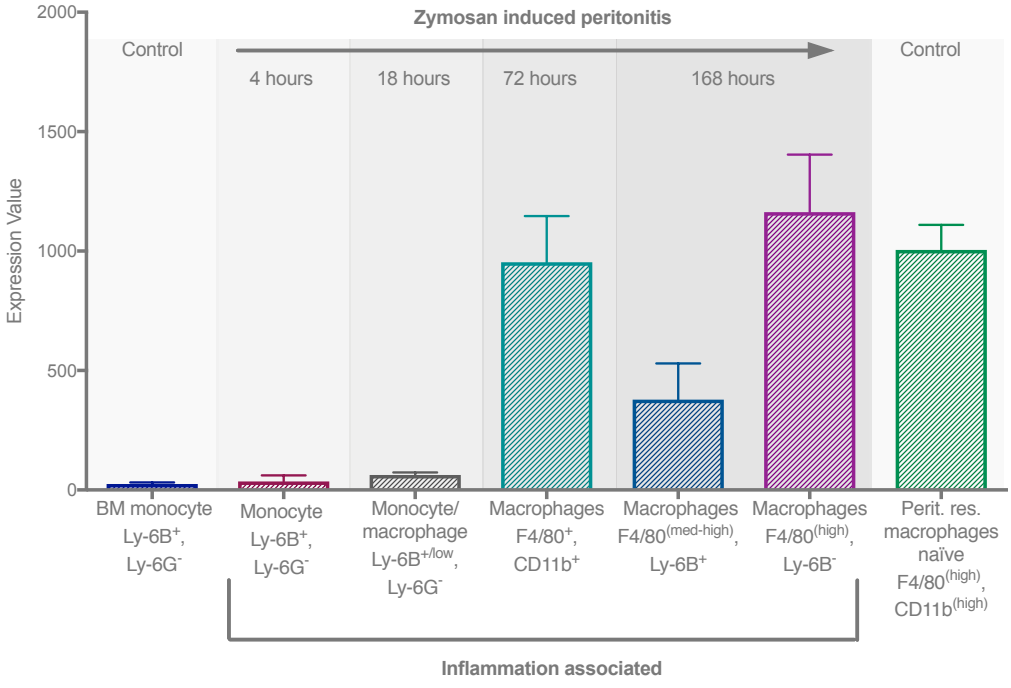


Figure 1.3 Selective expression of *Gata6* in pResMϕs.

Relative expression of *Gata6* in monocyte and Mϕ subsets during an inflammatory challenge with zymosan, at different time points. Data published in "Rosas, M. *et al.* The transcription factor GATA6 links tissue Mϕ phenotype and proliferative renewal. *Science* 344, 645–648 (2014)".

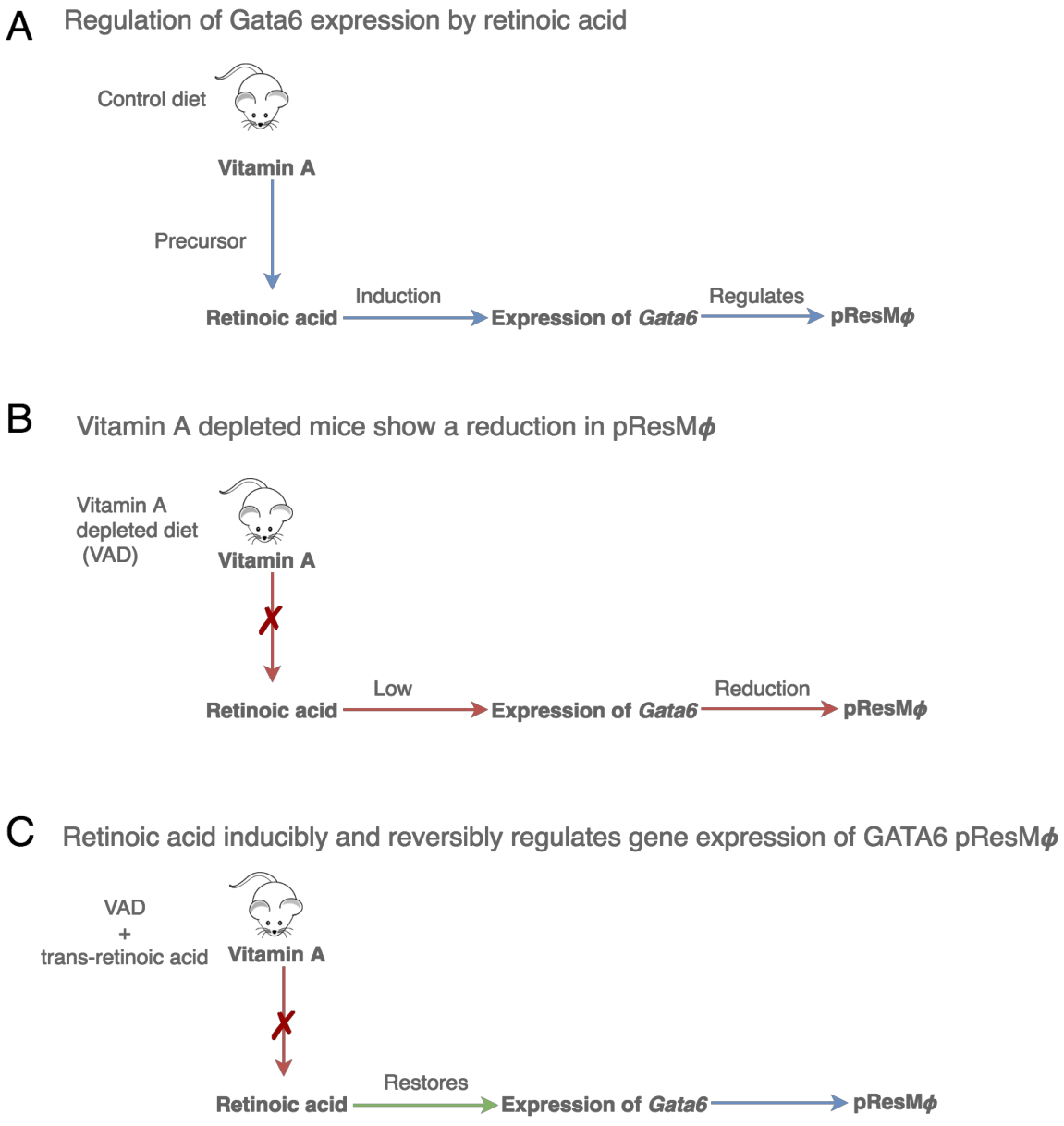


Figure 1.4 Tissue-specific signals control reversible program of localization and functional polarisation of Mϕ.

A – under normal conditions, Vitamin A acts as a precursor for retinoic acid. Which in turn, induces *Gata6* expression in pResMϕ; B – Vitamin A depleted mice have a reduced expression of *Gata6* and number of pResMϕ to when compared to control diet fed mice; C - Treatment with retinoic acid recovers *Gata6* expression.

1.1.2 GATA6 transcription factor

The GATA family is a group of evolutionarily conserved class IV zinc finger transcription factors, in vertebrates, that comprises six members, named GATA1 to -6, characterised by their ability to bind the DNA motif (G/A)GATA(A/T)[72]. All are related by a high degree of amino acid sequence homology within their zinc finger DNA-binding domains and protein-protein interaction with other transcriptional partners and/or cofactors[73]. The GATA family plays important roles in lineage specification and transdifferentiation, and was the first family of transcription factors shown to be able to substitute for octamer-binding transcription factor 4 (Oct4), a highly important pluripotency factor[74,75], and function as inducers of the reprograming process[76]. GATA proteins are classified into two subgroups: (i) GATA1/2/3 are expressed in hematopoietic cell lineages and are crucial for the proliferation of hematopoietic stem cells and differentiation of megakaryocytes and erythrocytes[72], (ii) GATA4/5/6 are largely found in tissues of mesodermal and endodermal origin, such as, heart, gut, and gonads where they play vital roles in controlling tissue-specific gene expression [76,77].

GATA6 is particularly associated with embryonic mesoderm/ endoderm[78], as well as mesoderm/ endoderm-derived organs, such as heart and gastrointestinal tract[79]. The *Gata6* mRNA encodes a single transcript containing two potential translational initiation sites. Translation at the first site encodes a long GATA6 isoform (L-type) that has a 146 amino acid residue extension in frame, while the second gives rise to a shorter isoform (S-type), with 78 amino acids[80] (Figure 1.5). GATA6 binds to the consensus sequence (A/T/C)GAT(A/T)(A) in the regulatory elements of target genes, which is very similar to the GATA motif (A/T)GATA(A/G). Additionally, GATA6 binds to a similar sequence containing the GATC but not GATG, indicating that it displays a broader sequence specificity for DNA binding than other members of the GATA family[81]. Even, though the *in vivo* roles of S- and L-type GATA6 have not yet been determined, *in vitro* reporter gene assays have demonstrated that the molecular activity of S-type GATA6 should be lower than that of L-type GATA6[81,82]. This difference has been attributed to the PEST sequence, found in L-type GATA6 which is usually linked

to protein degradation[83]. However, studies show that the PEST sequence in *Gata6* does not participate in protein degradation, but instead suggest a novel mechanism by which the PEST sequence contributes to additional DNA binding activity and the stable expression of L-type GATA6[81].

The *Gata6* gene is positioned upstream of other transcription factors, and its disruption in mice causes embryonic lethality between E6.5 and E7.5[82]. Moreover, GATA6 is suggested be a vital regulator of cell cycle progression[84]. This indicates that the expression level and modulation of GATA6 would be important for the phenotype of cells. Confirmed by evidence that GATA6 selectively regulates pResMϕ transcripts (described in section 1.1.1b)[51,57,71]. The cells display a failure to proliferate in states of inflammation[69] and enhanced apoptosis in the resting state[71], which in turn results in phenotypical changes of the pResMϕ population. This is supported by the reduction of pResMϕ in *Gata6* deficient mice[85], mentioned in section 1.1.1 b).

Importantly, the microarray analysis performed in 2014 by Rosas *et al.* in GATA6-deficient pResMϕ identified significant changes in 73 lipid metabolism related genes[69]. Among the first genes identified are *Abca1* (ATP-binding cassette, sub-family A (ABC1), member 1), *Alox15* (arachidonate 15-lipoxygenase), *Fabp4* (fatty acid binding protein 4) and *Ldlr* (Low Density Lipoprotein Receptor).

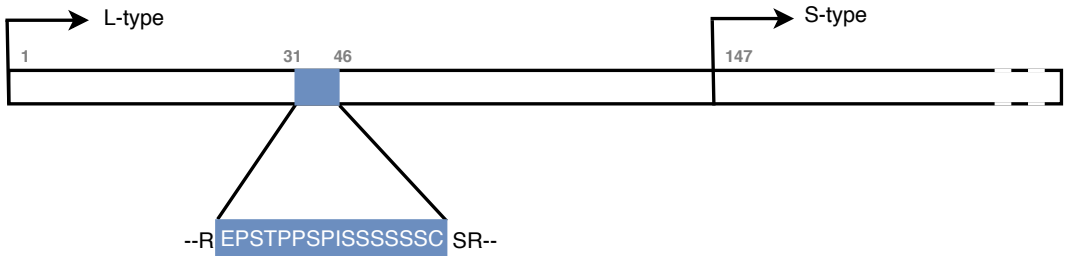


Figure 1.5 Schematic representation of GATA-6. L-type GATA-6 has an extension of 146 amino acid residues in front of S-type GATA-6.

Numericals indicate the amino acid residue numbers from the amino terminus of L-type GATA-6. PEST sequence (Glu^{31}–Cys^{46}).

- ABCA1 - mediates cholesterol and phospholipids efflux to lipid-poor apolipoproteins (apo-A1 and apoE), and mediates the transport of lipids between Golgi and cell membrane[86,87].
- ALOX15 - oxidises polyunsaturated fatty acids to their corresponding hydroperoxy derivatives, lipoxygenases are key players in regulation of the cellular redox homeostasis, which is an important element in gene expression regulation[88].
- FABP4 - fatty acid uptake, transport, and metabolism[88].
- LDLR - is involved in receptor-mediated endocytosis of specific ligands, apolipoproteins E and B-100[88].

The proteins encoded by these genes participate in different aspects of lipid metabolism, suggesting that alterations in their levels could potentially lead to a significant disruption of lipid levels.

Additionally, when looking into pResMϕ-specific genes, from two different sources (Gautier *et al.* 2012 and Rosas *et al.* 2014), it was verified that GATA6 regulates the expression of lipid metabolism related genes that are also responsible for pResMϕ phenotype (Figure 1.6). This indicated that *Gata6* might play an important part in the regulation of lipid metabolism in pResMϕ[69]. Thus, it is important to determine whether its deficiency in pResMϕ has a significant impact in lipid levels.

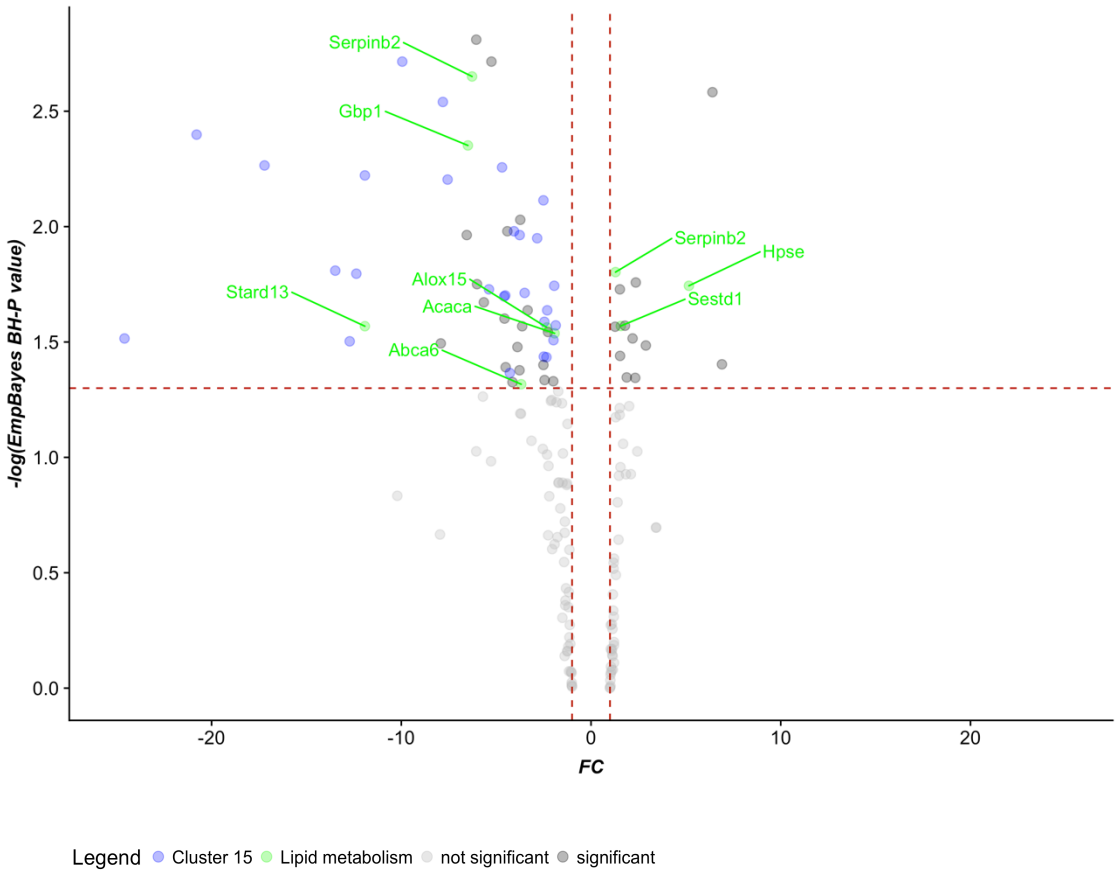


Figure 1.6 Alterations in peritoneal resident selective transcripts caused by GATA6 deficiency.

Examination of the genes identified by Gautier *et al.* and Rosas *et al.* as pResMϕ-specific. 'Volcano plot' of statistical significance against fold-change between *Gata6*-WT and -KO^{mye} pResMϕ, demonstrating the most significantly differentially expressed genes. Notably 2-fold down-regulated and up- regulated genes with p-value<0.05 are indicated; black- significantly regulated genes identified by Gautier *et al.* as pResMϕ-specific; blue- significantly regulated genes identified by Gautier *et al.* and confirmed by Rosas *et al.* as pResMϕ-specific (cluster 15); green- significantly regulated genes with functions related to lipid metabolism. A total of 67 genes with significantly altered expression levels were found (at least 2-fold altered and with p-value < 0.05), with 8 genes recognised as lipid metabolism-related genes. Statistical significance was assessed using the empirical Baysian Benjamini-Hochberg corrected p-value corresponding with that probeset. Data in this figure was obtained from Rosas *et al.* 2014.

1.2 Lipids and Mϕ phenotype

Lipids are important players in the regulation of cell phenotype and innate immune responses.

Herein, my goal is to establish a comprehensive picture of how GATA6 regulates lipid metabolism in murine peritoneal Mϕ. I will use novel top-down lipidomics techniques, for detection and analysis. I will first characterise the lipid pattern of *Gata6*-KO^{mye} versus WT pResMϕ using untargeted mass spectrometry (described in detail in the next section). I will then validate any changes of interest using targeted MS/MS. This data will then be combined with transcriptomics and functional data to create a picture of how lipid regulation contributes to the GATA6-dependent pResMϕ phenotype.

Ultimately, creating predictive models by developing/using comprehensive novel lipidomics methodologies requires a degree of mechanistic understanding of lipid metabolism. Nevertheless, it is a major goal from a systems biology standpoint because it will aid in the understanding of Mϕ biology that, in turn, may result in therapeutic outcomes.

1.2.1 Lipidomics

The "omics" revolution has afforded an invaluable understanding of genes, mRNAs, metabolites and proteins of mammalian cells, biological systems and disease[89–91]. Within this, lipidomics describes the analysis of the lipidome, providing comprehensive description of all lipids in a sample. Lipidomics involves systems-level identification and quantitation/comparison of lipid molecular species [92–96]. Several technologies have been used to identify, quantify and understand the molecular structure of lipids. However, it is clear that the advent of modern mass spectrometry (*e.g.*, electrospray ionization (ESI) and matrix-assisted laser desorption/ionization) helped drive the progress of lipidomics[97]. This made possible the systematic analysis of global lipidomes of mammalian cells and tissues, as well as the changes in the lipids associated with cellular function and disease[93,95,96,98].

The International Lipid Classification and the Nomenclature Committee, together with the Lipid Metabolites and Pathways Strategy (LIPID MAPS) Consortium, described eight lipid categories, divided into classes and subclasses, based on their functional backbones and biochemical principles[99] (Table 1.2). Nevertheless, a comprehensive database of lipid structures is not enough to fully understand their multiple biological roles in cell biology and pathology. In order to clarify lipid functions and the enzymes related to their metabolism, it is essential to organize them in the context of biological pathways and derive their associations and interactions with enzymes and other molecules[100]. Lipid metabolism related pathways control the production and degradation of major lipid categories. Even though some of these pathways have been studied in detail, it is still unclear how they interact in homeostatic conditions[101]. Nonetheless, advances in mass spectrometry have made it possible to qualitatively and quantitatively analyse lipids[102–105]. Currently, Liquid chromatography/mass spectrometry (LC-MS) is the best analytical technique to separate, identify, and quantitate lipids[92], and together with transcriptional approaches, is helping bridge the gap in our knowledge of how different lipid pathways are integrated.

Over the past two decades, the LIPID MAPS consortium took on the challenge of establishing core facilities for each of the key functional classes of lipids, in an effort to identify and quantify changes in metabolites by developing new techniques for lipid detection and analysis, as well as systematically analysing the vast amount of data generated[92]. The consortium initially focussed on Mϕ-like cells, due to the fact that these were already widely studied in connection to lipid metabolism. However, their analytical strategy comprised the use of gold standard quantitative MRM methods, that measure only known lipids and therefore do not provide information regarding lipids that have not been previously identified.

The study of the lipidomic response of the murine Mϕ RAW cell line to Kdo2-lipid A was the first comprehensive systems-level quantitative analysis of a mammalian cell lipidome[95]. This is of particular relevance because it provides information on the lipid species I can expect to find in Mϕ-like cells. Additionally, I will perform a lipidomic comparison between RAW and pResMϕ cells which will help establish the similarities/differences among the two lipidomes under basal conditions.

Table 1.2 Lipid Categories and classes[106]**.**

Lipid Category	Lipid Class
Fatty Acyls [FA]	Fatty Acids and Conjugates [FA01]
	Octadecanoids [FA02]
	Eicosanoids [FA03]
	Docosanoids [FA04]
	Fatty alcohols [FA05]
	Fatty aldehydes [FA06]
	Fatty esters [FA07]
	Fatty amides [FA08]
	Fatty nitriles [FA09]
	Fatty ethers [FA10]
	Hydrocarbons [FA11]
	Oxygenated hydrocarbons [FA12]
	Fatty acyl glycosides [FA13]
	Other Fatty Acyls [FA00]
Glycerolipids [GL]	Monoradylglycerols [GL01]
	Diradylglycerols [GL02]
	Triradylglycerols [GL03]
	Glycosylmonoradylglycerols [GL04]
	Glycosyldiradylglycerols [GL05]
	Other Glycerolipids [GL00]
Glycerophospholipids [GPL]	Glycerophosphocholines [GP01]
	Glycerophosphoethanolamines [GP02]
	Glycerophosphoserines [GP03]
	Glycerophosphoglycerols [GP04]
	Glycerophosphoglycerophosphates [GP05]
	Glycerophosphoinositols [GP06]
	Glycerophosphoinositol monophosphates [GP07]
	Glycerophosphoinositol bisphosphates [GP08]
	Glycerophosphoinositol trisphosphates [GP09]
	Glycerophosphates [GP10]
	Glyceropyrophosphates [GP11]
	Glycerophosphoglycerophosphoglycerols [GP12]
	CDP-Glycerols [GP13]
	Glycosylglycerophospholipids [GP14]
	Glycerophosphoinositolglycans [GP15]
	Glycerophosphonocholines [GP16]
	Glycerophosphonoethanolamines [GP17]
	Di-glycerol tetraether phospholipids (caldarchaeols) [GP18]
	Glycerol-nonitol tetraether phospholipids [GP19]
	Oxidized glycerophospholipids [GP20]
	Other Glycerophospholipids [GP00]
Sphingolipids [SL]	Sphingoid bases [SP01]
	Ceramides [SP02]
	Phosphosphingolipids [SP03]
	Phosphonosphingolipids [SP04]
	Neutral glycosphingolipids [SP05]
	Acidic glycosphingolipids [SP06]
	Basic glycosphingolipids [SP07]
	Amphoteric glycosphingolipids [SP08]
	Arsenosphingolipids [SP09]
	Other Sphingolipids [SP00]
Sterol Lipids [ST]	Sterols [ST01]
	Steroids [ST02]
	Secosteroids [ST03]

	Bile acids and derivatives [ST04]
	Steroid conjugates [ST05]
	Other Sterol lipids [ST00]
Prenol Lipids [PR]	Isoprenoids [PR01]
	Quinones and hydroquinones [PR02]
	Polyprenols [PR03]
	Hopanoids [PR04]
	Other Prenol lipids [PR00]
Saccharolipids [SL]	Acylaminosugars [SL01]
	Acylaminosugar glycans [SL02]
	Acyltrehaloses [SL03]
	Acyltrehalose glycans [SL04]
	Other acyl sugars [SL05]
	Other Saccharolipids [SL00]
Polyketides [PK]	Linear polyketides [PK01]
	Halogenated acetogenins [PK02]
	Annonaceae acetogenins [PK03]
	Macrolides and lactone polyketides [PK04]
	Ansamycins and related polyketides [PK05]
	Polyenes [PK06]
	Linear tetracyclines [PK07]
	Angucyclines [PK08]
	Polyether antibiotics [PK09]
	Aflatoxins and related substances [PK10]
	Cytochalasins [PK11]
	Flavonoids [PK12]
	Aromatic polyketides [PK13]
	Non-ribosomal peptide/polyketide hybrids [PK14]
	Phenolic lipids [PK15]
	Other Polyketides [PK00]

1.2.2 Bioactive lipids

The concept of bioactive lipids originated from the recognition that lipids are not only involved in cell membrane structure and energy storage[107]. Indeed, some lipids have a plethora of roles, such as, immune regulation, inflammation and maintenance of homeostasis[108]. Accordingly, variations in the levels of such lipids leads to pathophysiological consequences[107]. Studies over the past four decades have defined several lipid families as key inter- and intracellular lipid signalling molecules[109–112].

Bioactive lipids, can be divided into four main families according to their biochemical functions (Figure 1.7), *i.e.*, classical eicosanoids, are primarily involved in mediating or resolving inflammatory responses[113], lyso-glycerophospholipids/SL and endocannabinoids (eCBs), are generated from ω-6 or ω-3 essential polyunsaturated fatty acids (PUFA) precursors, that are esterifed into membrane lipids and act by binding to and activating specific G protein-coupled receptors (GPRs), while phospholipids are involved in metabolism and cell signalling[108]. Sphingolipids (SL), on the other hand, have key roles in cell migration, proliferation, differentiation, enhancing/dampening immune function and apoptosis[114].

The first bioactive SL identified was sphingosine (So), which exerts pleiotropic effects on protein kinases. So and its related sphingoid bases regulate endocytosis, the actin cytoskeleton, cell cycle and apoptosis (Figure 1.8)[115]. Since its discovery, even more emphasis has been placed on other bioactive SL such as ceramide (Cer) and sphingosine-1-phosphate (S1P). Cer mediates cell-stress responses, such as cell senescence[116] and apoptosis[117], whilst S1P is crucial for cell survival, cell migration and inflammation[118]. Other additions to the bioactive SL family include ceramide-1-phosphate (C1P, has roles in inflammation and vesicular trafficking[119,120]), and glucosylceramide (GlcCer) (involved in post-Golgi trafficking and in drug resistance[120]).

Recently, SL have emerged as signalling molecules that regulate many cell functions, including in the regulation of inflammatory responses[114]. In the following Chapters I will perform a comprehensive analysis of the pResMϕ

lipidome, where I demonstrate that SL are significantly altered in GATA6 deficient cells. Thus, due to their relevance in Mϕ metabolism, and to this project, the SL category will be further introduced in the next section.

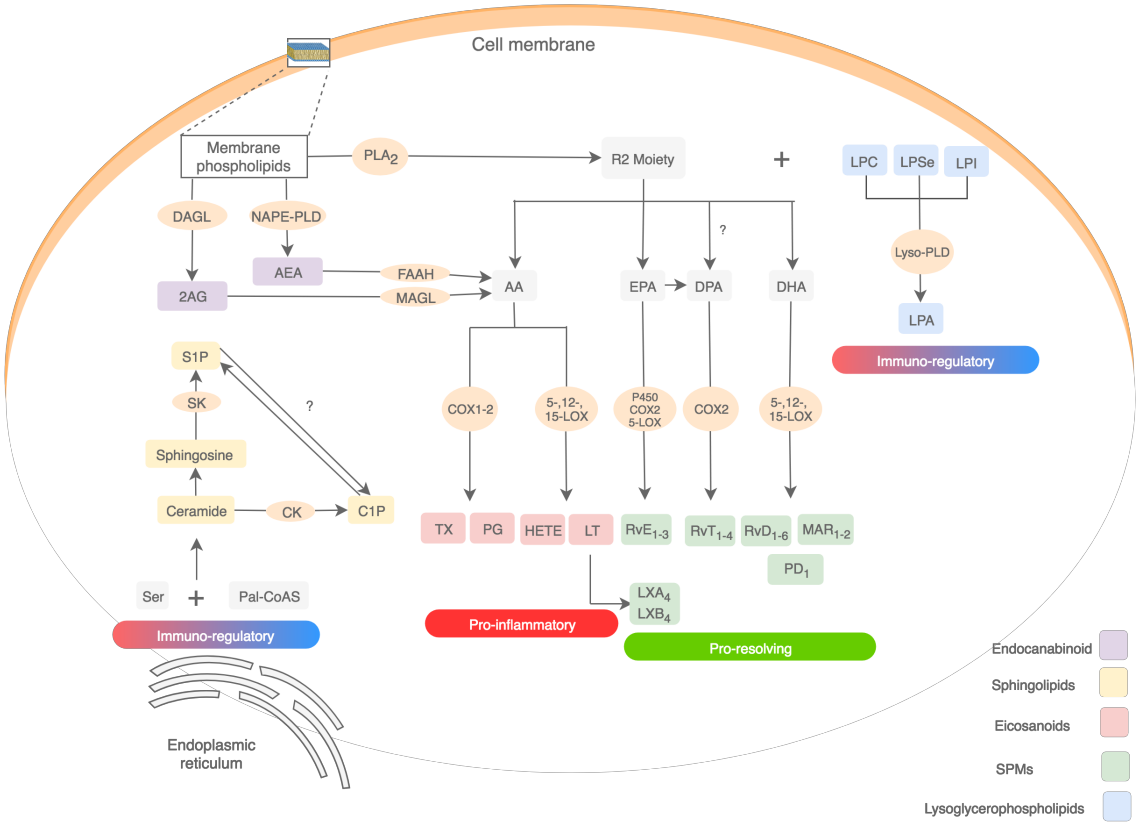


Figure 1.7 Metabolic pathways of key families of bioactive lipids.

2-AG, 2 arachidonoylglycerol; AA, arachidonic acid; AEA, arachidonoylethanolamide; C1P, ceramide-1-phosphate; CK, ceramide kinase; COX, cyclooxygenase; DAGL, diacylglycerol lipase; DHA, docosahexaenoic acid; DPA, docosapentaenoic acid; EPA, eicosapentaenoic acid; FAAH, fatty acid amide hydrolase; HETEs, hydroxyeicosatetraenoic acids; LOX, lipoxygenase; LPA, lysophosphatidic acid; LPC, lysophosphatidilcholine; LPI, lysophosphatidylinositol; LPSer, lysophosphatidylserine; LTs, leukotrienes; LX, lipoxin; Lyso-PLD, lyso-phospholipase D; MAGL, monoacylglycerol lipase; MaR, maresin; NAPE-PLD, N-arachidonoylphosphatidylethanolamide-specific phospholipase D; Pal-CoASH, palmitoyl coenzyme A; PD, protectin; PDX, protectin DX; PGs, prostaglandins; PLA2, phospholipase A2 ; Rv, resolvin; S1P, sphingosine-1-phosphate; SK, sphingosine kinase; TXs, thromboxanes. (adapted from Chiurchiù, V., Leuti, A. & Maccarrone, M. Bioactive lipids and chronic inflammation: Managing the fire within. Front. Immunol. 9, (2018)).

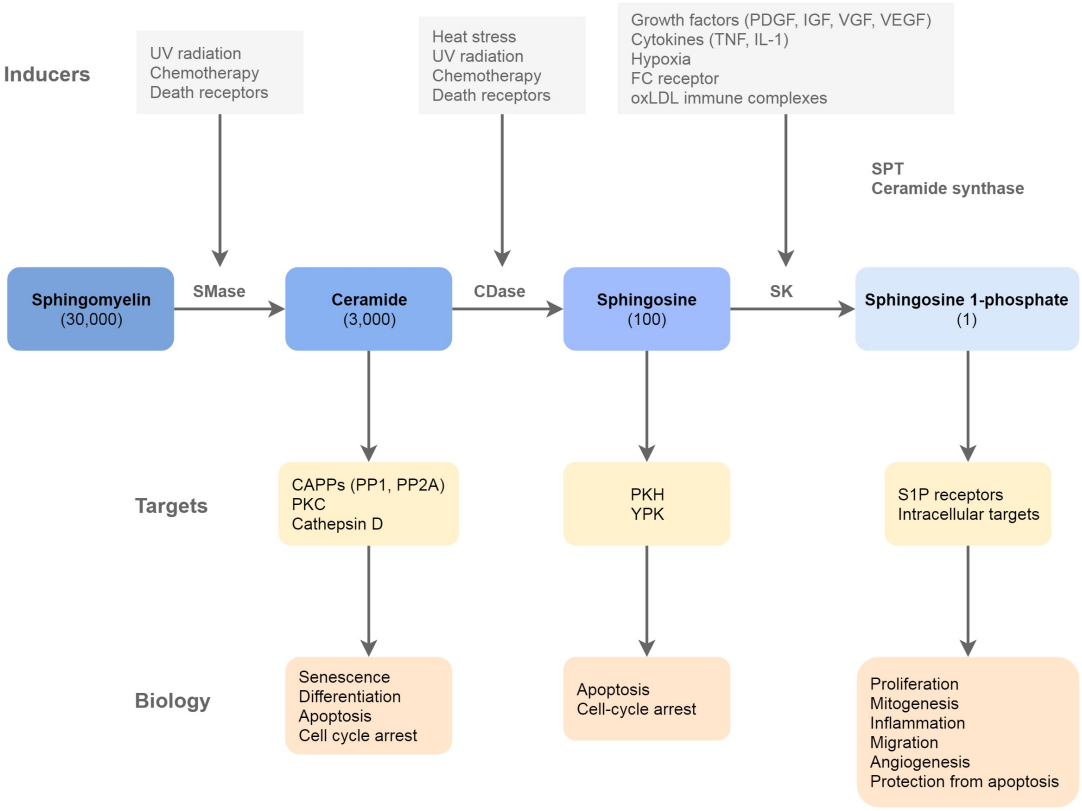


Figure 1.8 Roles of SL in biology. The figure depicts the potential participation of the bioactive SL, Cer, So and sphingosine-1phosphate (S1P) in cell biological responses.

Ceramide is at the centre of SL biology, it can be produced by the breakdown of sphingomyelin by sphingomyelinases (SMases) or synthesized *de novo* by serine palmitoyl transferase (SPT) and Cer synthase. These reactions can be induced by diverse stimuli (grey boxes). So and S1P can be generated by ceramidases (CDases) and So kinases (SKs). Which in turn interact with specific protein targets such as phosphatases, kinases and G protein coupled receptors (S1P receptors). The numbers in brackets indicate the comparative levels of these SL. CAPP, Cer activated Ser–Thr phosphatase; IGF, insulin-like growth factor; IL-1, interleukin-1; oxLDL, oxidized low-density lipoprotein; PDGF; platelet-derived growth factor; PKC, protein kinase C; PKH, PKB homologue; TNFa, tumour necrosis factor-a; VEGF, vascular endothelial growth factor; YPK, yeast protein kinase. Adapted from Hannun, Y. A. & Obeid, L. M. Principles of bioactive lipid signalling: lessons from SL. Nat. Rev. Mol. Cell Biol. 9, 139–150 (2008).

1.2.3 Sphingolipids (SL)

SL are a category of bioactive lipids where the defining component is a long-chain or sphingoid base. This long chain, usually a So, is the basic element and the simplest possible functional SL (Figure 1.9). The first SL was described by German biochemist J. L. W. Thudichum in 1884. He named it sphingomyelin (SM) after the sphinx in Greek mythology (for the many enigmas presented by its chemical structure), isolated from myelin[121]. However, SM is not restricted to the brain, it is one of the most abundant SL[122], it forms a concentration gradient along the secretory pathway with the highest concentration in the plasma membrane (where it accumulates in the exoplasmic leaflet)[123].

Ceramides, which contain a fatty acid linked by an amide bond to the sphingoid base, function both as a key players in cell signalling and as the precursors of more complex SL with an immense range of functions in tissues[124]. Complex SL may be formed by addition of head groups to Cer (Figure 1.9):

- Sphingomyelins can have either a phosphocholine or phosphoethanolamine group attached with an ester linkage to the 1-hydroxy group of a Cer.
- Glycosphingolipids (GSL) are Cer molecules with one or more sugar residues joined with a glycosidic linkage at the 1-hydroxyl position.
    - Cerebrosides are Cer molecules with a single glucose or galactose at the 1-hydroxy position.
        - Sulfatides are cerebrosides with a sulfate group attached.
    - Gangliosides have at least three sugars, one of which must be sialic acid.

Within membranes, SL are associated with cholesterol, and together they help form lipid domains such as caveolae and rafts[125]. Their metabolism is controlled

at multiple levels, through the regulation of enzyme expression, post-translational alterations and allosteric mechanisms[126]. Some of these are specific to certain cell types, either regulating which SL are synthesized during different stages of cell development, or in response to specific signals.

At least five different long-chain bases are known in mammalian cells. Over 20 species of fatty acid (with varying chain lengths, degrees of saturation and hydroxylation) can be attached to the sphingoid base, and around 500 different carbohydrate structures have been described in GSL. Due to the combinatorial nature of SL biosynthesis, these backbone, headgroup, and acyl chain variations give rise to tens of thousands of mammalian SL species and this complexity demands highly organized and intricate mechanisms of regulation both at cellular and biochemical levels, which still have not been fully characterised[124,127].

1.2.4 The interconnectivity of SL metabolism

As mentioned in the previous section, SL metabolism can be regulated at numerous stages, from control of enzyme expression, to post-translational modifications and allosteric mechanisms[124]. Therefore, the coordinate production and removal of SL species requires a delicate balance between *de novo* biosynthesis, turnover, and recycling[128](Figure 1.10).

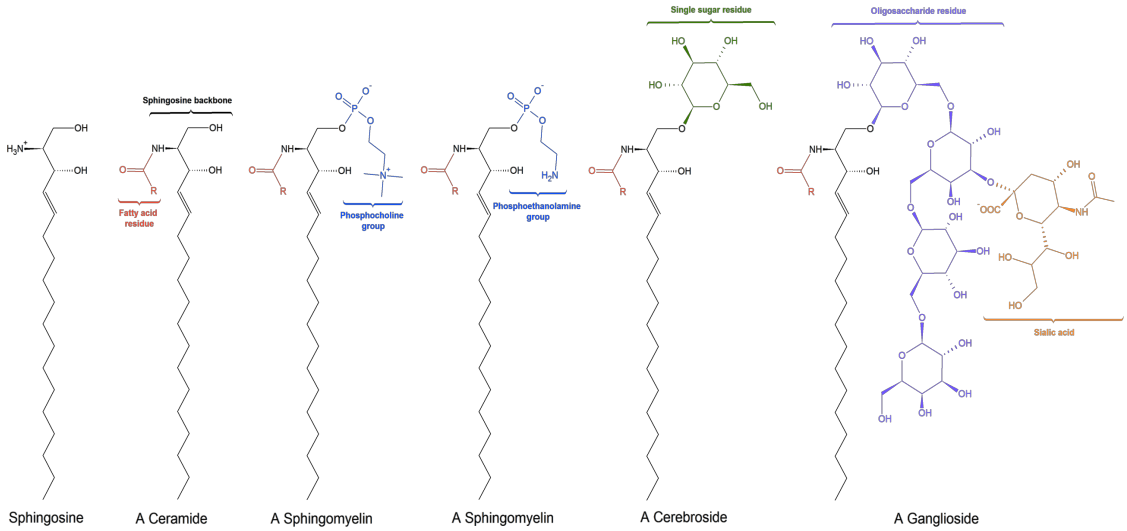


Figure 1.9 SL classes.

SL are a lipid category, where the lipid backbone is a sphingoid base (sphingosine (So) or sphinganine (Sa)). Modifications of this basic structure consist in the addition of an amide-linked fatty acid or sugar residues that lead to the formation of more complex lipid species such as ceramide (Cer), sphingomyelin (SM), cerebroside or ganglioside.

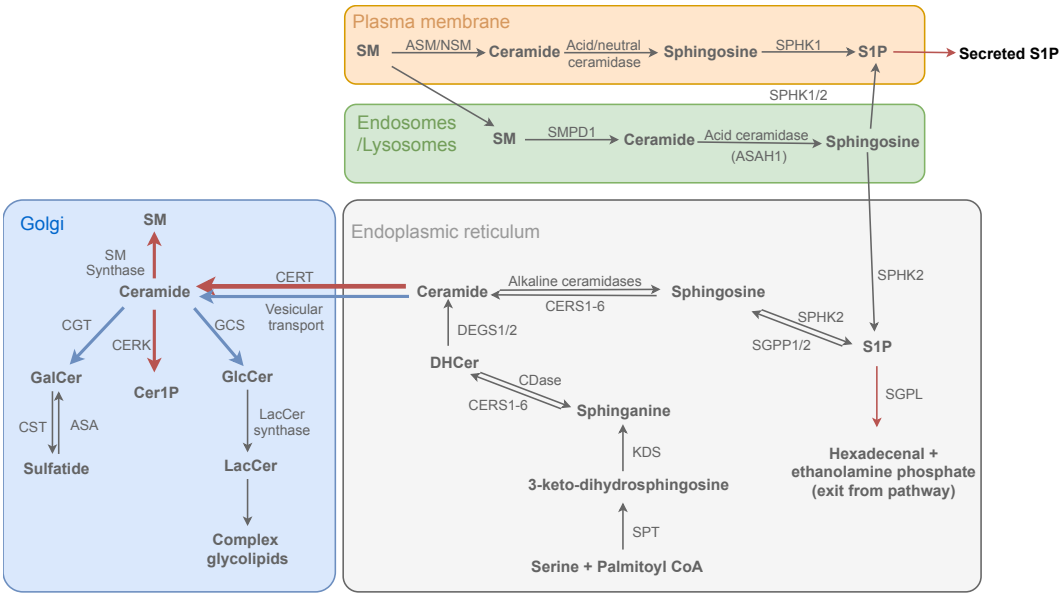


Figure 1.10 SL metabolism and basic structural units of SL biosynthesis.

Ceramide can be synthesized *de novo* from l-serine and palmitoyl-CoA (Endoplasmic reticulum), through hydrolysis of sphingomyelin (plasma membrane and Endosomes/lysosomes) or through hydrolysis of GSL and sulfatides (Golgi). ASA: Arylsulfatase A; ASAH1: Acid ceramidase; SMPD1: Acid sphingomyelinase; Cer1P: Ceramide 1-phosphate; CERK: Ceramide kinase; CERS: Ceramide synthase; CERT: Ceramide transfer protein; CGT: Ceramide galactosyltransferase; CST: Cerebroside sulfotransferase; DEGS1/2: Dihydroceramide desaturase 1 or 2; GCS: Glucosylceramide synthase; KDS: 3-ketosphinganine reductase; LacCer synthase: Lactosylceramide synthase; NSM: Neutral sphingomyelinase; S1P: Sphingosine 1-phosphate; SGPL: S1P lyase; SGPP1/2: Sphingosine 1-phosphate phosphatase 1 or 2; SM: Sphingomyelin; SPHK1/2 : Sphingosine kinases 1 or 2; SPT: Serine palmitoyltransferase.

a) De novo SL biosynthesis

In mammalian cells, the SL metabolic pathway (Figure 1.10) begins with the *de novo* pathway, in the endoplasmic reticulum (ER)[129], where four enzyme groups cooperate to form Cer of different acyl chain lengths from non-SL precursors[130]. The enzyme complex serine palmitoyltransferase (SPT) catalyses the first step of SL biosynthesis, the condensation of L-serine and palmitoyl-CoA to form 3-ketosphinganine (3KSa)[128,131,132], followed by the rapid reduction of 3KSa to form the sphingoid base dihydrosphingosine (*a.k.a.* sphinganine) (Sa) a by 3-ketosphinganine reductase (3KSR)[129,133,134]. Then, Sa is phosphorylated by So kinase (SphK) to form sphinganine 1- phosphate (Sa1P) or variable length fatty acid chain is transferred to the free amine group of Sa by a family of six (dihydro)Cer synthases (CerS)[135] to form dihydroceramide (DH-Cer), each of these enzymes utilizes fatty acyl-CoA substrates of relatively distinct chain lengths resulting in the formation of distinct DH-Cer subspecies[128,130]. After its generation, dihydroceramide is metabolised by a desaturase (DES1 and 2) which introduces a double bond, thus converting DH-Cer into Cer with various chain lengths. These Cers will then act as precursors to complex SL[128,129,134,136,137].

This initial phase of SL synthesis is where complex SL emerge, through differential addition of acyl chains to the Sa/So backbone by action of CerS[124]. Cer is a membrane bound molecule, with very low solubility in aqueous environment, its hydrophobic nature prevents spontaneous relocation through the cytosol[138], therefore it must travel from one membrane to another by facilitated mechanisms[118]. The cell possesses two main mechanisms for Cer transport, vesicular transport or through the Cer transfer protein (CERT)[138]. Additionally, Cer transported to the Golgi by CERT is preferentially incorporated into SM, whilst vesicular transport is thought to be the major pathway responsible for conveying Cer to the cis-Golgi for GSL synthesis[123].

b) Turnover pathway

(DH)Cer can also be produced by the breakdown or turnover of complex SL through the action of specific hydrolases, *i.e.* sphingomyelinases, glycohydrolases and lipid phosphatases (Figure 1.10)[134]. Furthermore, (DH)Cer can also be

metabolized by ceramidases that catalyse the hydrolysis of (DH)Cer into a free fatty acid and a sphingoid base, Sa or So[134,139,140]. These sphingoid backbones, can then be recycled via reacylation for entry back into the SL pathway[134,139] or it can be phosphorylated by SphK, producing Sa1P or So1P[134,141]. Conversely, Sa1P/So1P can be de-phosphorilated through the action of phosphatases or degraded by sphingosine 1-phosphate lyase (lyase) to form ethanolamine phosphate and hexadec(an/en)al[134,142], which can then be reduced to palmitate and re-incorporated into SL biosynthesis[134].

SL biochemical interconnections allow cells to organise cellular responses by regulating SL interconversions. For instance, activation of SMase leads to the conversion of SM into Cer. Conversely, the subsequent action of ceramidase, Cer kinase, SM synthase or GlcCer synthase converts the Cer signal to one that is mediated by either So, Cer1P, DAG or GlcCer (Figure 1.11)[118]. In the majority of cell types, the concentration of SM is one order of magnitude higher than that of Cer, hence, even the smallest changes in SM may result in significant changes in Cer amounts. Additionally, more than 50 distinct molecular species can be classified as Cer, and they may have their own separate metabolic networks, resulting in an extremely high level of complexity in SL signalling[130].

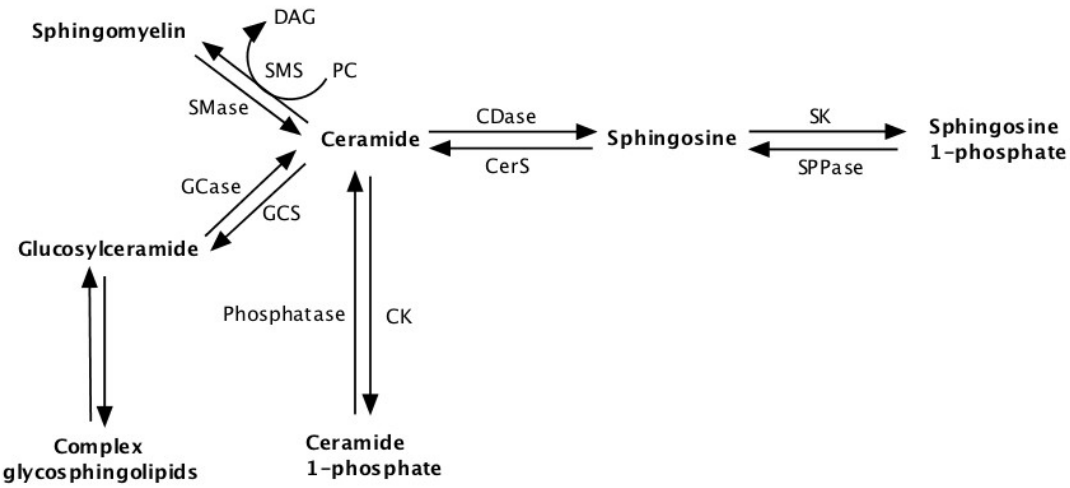


Figure 1.11 SL metabolic pathway.

Ceramide is at the core of SL metabolism, it can be synthesised by the action of desaturases, and then converted into other bioactive lipid species. Herein, only a pathway for a generic Cer is depicted. CDase, ceramidase; CK, Cer kinase; DAG, diacylglycerol; GCase, glucosyl ceramidase; GCS, glucosylceramide synthase; PC, phosphatidylcholine; SK, sphingosine kinase; SMase, sphingomyelinase; SMS, sphingomyelin synthase; SPPase, sphingosine phosphate phosphatase; SPT, serine palmitoyl transferase.

1.2.6 Liquid Chromatography–Mass Spectrometry

The analysis of lipids advanced greatly, when liquid chromatography (LC) was combined with mass spectrometry (MS)[143]. MS is an important analytical technique that measures the mass-to-charge ratio (*m/z*) of ions to identify and quantify molecules in samples with various degrees of complexity. However, it was the advances in MS, observed in the past few decades that propelled the progress in the field of lipid research, in particular, soft ionization techniques such as ESI and MALDI.

A mass spectrometer usually consists of an ion source, a mass analyser, a detector and a data processing system (Figure 1.12). The current variety of combinations of ionisation, mass analysing, and detection methods, make MS the technique of choice for lipid analysis[144]. The ability to ionise non-volatile molecules and allow their entry into analysers has revolutionised lipid research, through enabling the determination of molecular weight, elemental composition, abundance, and structure [97].

Though there are several instrument configurations available, MS system operation can be organized into three main divisions (Figure 1.12):

(1) generation of ions;

(2) separation of ions based on mass and charge in a mass analyser;

(3) detection of ions and instrument output.

Depending upon the type of ionization used, these steps fully or partially occur under vacuum pressure to drive ion movement forward through the instrument[145].

The ionisation method determines the classes of substances that can be analysed[144]. For an MS-based analysis, molecules must be ionized, or electrically charged, to produce individual ions. MS instruments analyse molecules by relating the mass of each molecule to its charge. This identifying characteristic is specific to each molecule and is referred to as *m/z*. Thus, if the molecule has a

single charge ($z = 1$), the m/z ratio will be equal to the molecular mass of the ionized species[146].

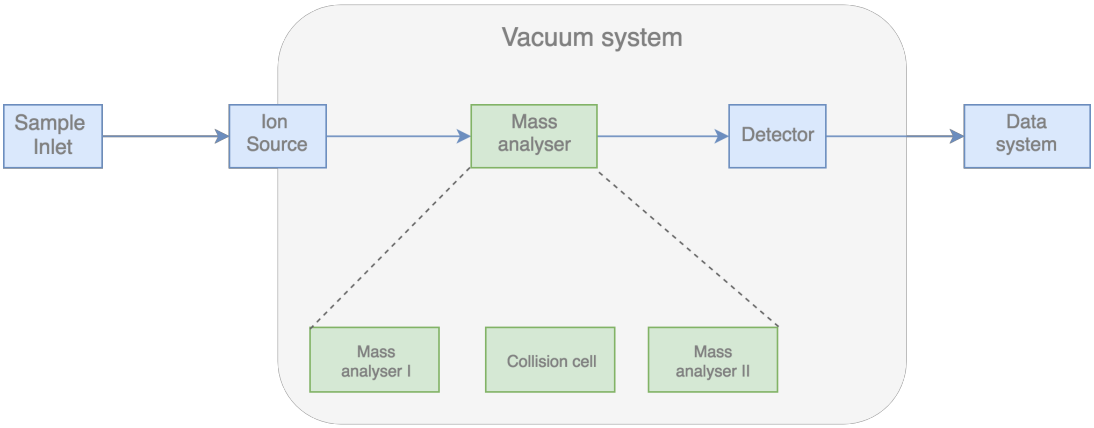


Figure 1.12 Basic components of a triple quadrupole Mass Spectrometer.

Sample inlet – introduction of samples in the ion source, while maintaining high vacuum requirements; Ion source – production of gaseous ions from the substance being studied; Mass analyser – for resolving the ions into their characteristics mass components according to their *m/z*.; Detector – detection of ions and recording the relative abundance of each of the resolved species.; Data system – data analysis and interpretation.

The combination of mass analyser and detector determines the quality and reliability of analysis[144]. Mass analysers can be divided into several generic types, quadrupole, magnetic sector, ion trap, time-of-flight (TOF), or Fourier transform (FT). Depending on the combination, they can be used for analysis of both analytes and their fragments[147]; *e.g.* triple quadrupole and quadrupole/TOF hybrids. Furthermore, the same analyser can perform both MS and MS^n(multi-stage mass spectrometry) [144].

The analytical power of a mass spectrometer lies in its resolution, its ability to discern one molecular mass from another[148]. High-resolution instruments possess increased mass accuracy. Defined as the mass difference that can be detected by the analyser divided by the observed, or true mass, in parts per million[149].

Equation 1.1

$$Mass\ accuracy\ (\Delta ppm) = \frac{True\ mass - Observed\ mass}{Observed\ mass} \times 10^6$$

Due to the great structural diversity in lipids, there is no single methodological platform currently available that can perform the full characterisation of the lipidome. Instead, there are several strategies, tailored to different lipid categories/classes. Thus, it is necessary to combine complementary approaches in order to achieve high quality datasets. The two approaches used in the course of this project will be further discussed in the following sections.

a) Untargeted LC-MS - Orbitrap Elite

An untargeted MS approach uses 'global profiling' methods to detect lipid species, *i.e.*, MS scans are performed over a wide *m/z* value, allowing the detection of lipids purely based on their *m/z*. Herein, I opted for an untargeted LC-MS approach using a high-resolution (up to 240 000 at *m/z* 400 for a 768 ms transient[144]) mass spectrometer, belonging to the FT family, the Orbitrap Elite. This instrument measures a frequency that is then converted to *m/z* via a Fourier

transform[150]. The Orbitrap Elite has the ability to differentiate ions with closely located *m/z*[151]. However, despite this high accuracy, with this untargeted approach I can only provide putative annotations for the ions. No distinction can be made between structurally different lipids of the same mass (isobars), unless LC-MS is used, supplying retention time information. In my experiments, I will use LC-MS on an Orbitrap platform and then analyse my data using multivariate data analysis, in order to determine which lipid categories should be further studied using targeted approaches.

b) Targeted lipidomics – Tandem MS

In order to perform a targeted lipidomic analysis it is necessary to use a mass spectrometer possessing multiple mass analysers[152]. Herein I will focus on the instrument selected for this project, a triple quadrupole mass spectrometer (QTrap 6500). This instrument can be used in single scan MS to provide information about precursor masses (similar to untargeted analysis), however it is more commonly used in tandem scan mode (MS/MS), providing additional information regarding specific lipid species[153]. This mode is more selective and sensitive because it uses specific parent-to-daughter transitions. Additionally, when coupled with LC it also supplies retention time information, that can further be used to complement the lipid identification. The main MS/MS modes are described in Figure 1.13, product-ion analysis, neutral-loss scan (NLS), precursor-ion scan (PIS), and selected reaction monitoring (SRM)[154]. SRM is monitoring only a single fixed mass window. When multiple transitions are monitored during a chromatographic run, the data acquisition method is called multiple-reaction monitoring (MRM). Thus, MRM is the application of SRM to multiple product ions from one or more precursor ions, which allows for the quantitation of multiple peaks from a single chromatographic run allowing for simultaneous detection and quantitation of many analytes (multiplexing)[146]. When a precursor ion enters the MS, it is selected in Q1 by its *m/z* ratio and it is transferred to the collision cell (Q2) where it is fragmented. The fragments obtained are monitored in Q3 and the spectrum corresponding to the *m/z* of interest is recorded[153]. Ions with different *m/z* ratios will be filtered out and will not reach the detector.

This type of scan is often used for identification and quantitation of lipids, thus providing complementary information, to that obtained using an untargeted approach. Thus, the combination of the two methodologies facilitates a more powerful lipidomic analysis.

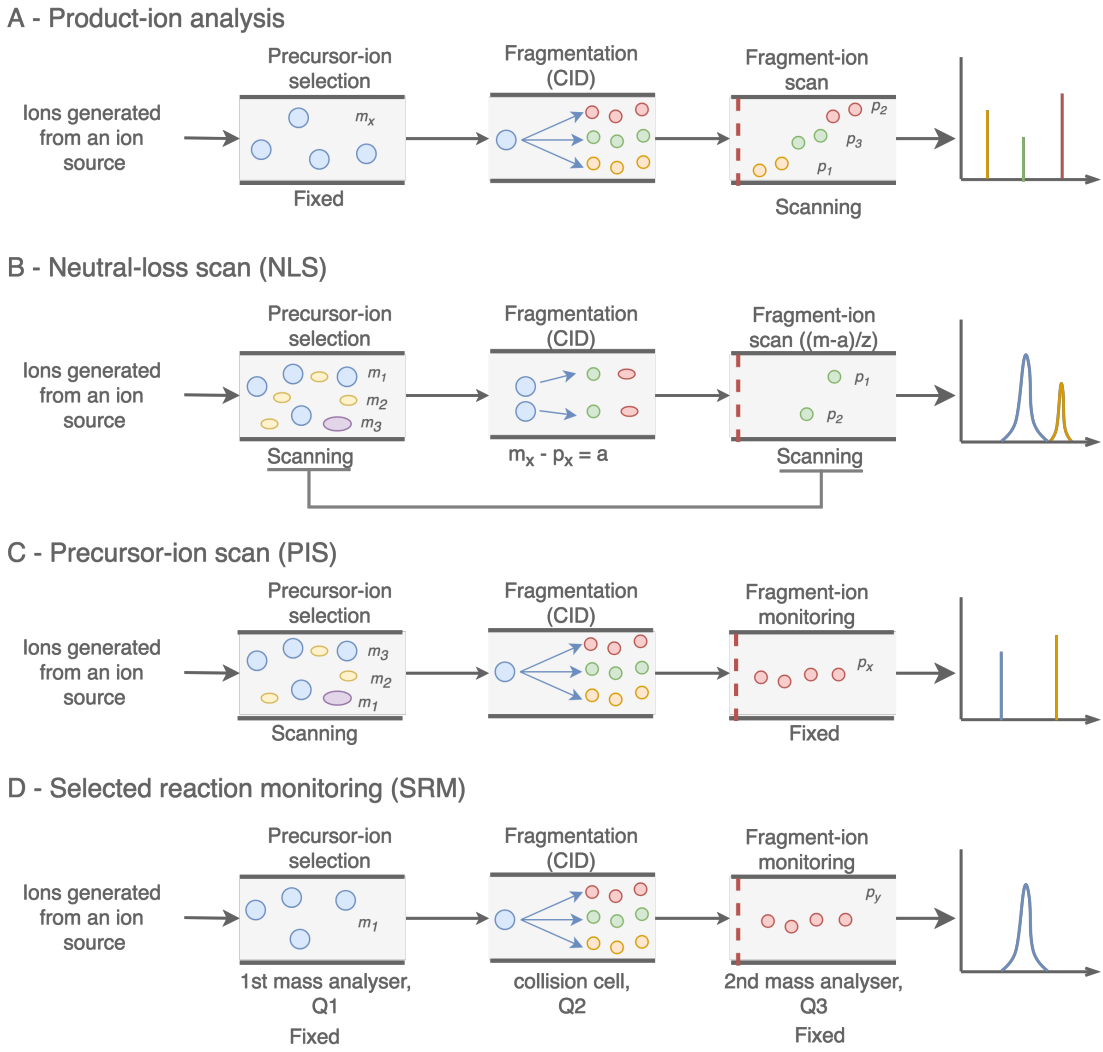


Figure 1.13 Tandem mass spectrometry main scan modes of analysis in a triple quadrupole mass spectrometer.

CID stands for Collision induced dissociation, and symbolism for representation of the various scan modes. (A) Product ion scan: select Q1 precursor ion and scan Q3 product ion. (B) Neutral loss scan: scan all ions in Q1 and select ions with neutral loss in Q3. (C) Precursor ion scan: scan precursor ion in Q1 and select certain fragment ion in Q3, all collision induced dissociation carried out in Q2. (D) Selected reaction monitoring: select precursor ion in Q1 and monitor one or more fragment ions in Q3.

1.3 General Hypothesis

As described earlier, prior to this study, GATA6 was recognised as a master regulator of pResMϕ phenotype and the expression of lipid metabolism-related genes was found to be significantly altered in murine GATA6 deficient pResMϕ.

Herein I proposed that, myeloid deficiency of GATA6 (*Gata6*-KO^{mye} mice) is associated with significant changes in lipid metabolism, at a transcriptional and molecular level.

In order to prove my hypothesis, it is crucial to undertake a detailed analysis of lipid metabolism in GATA6 deficient pResMϕ. This will include studies on the following:

- Generation of appropriate targeted/untargeted mass spectrometry workflows
- Canonical pathways
- Gene networks
- Global lipid levels
- Significantly altered lipid categories
- Integration of transcriptional and lipidomics data
- Quantitation of significantly altered lipids

Chapter 2 Materials and Methods

2.1 Buffers, solutions and Media

Unless otherwise stated, chemicals were from Merck chemicals Ltd (Southampton, UK) or Thermo Fisher Scientific Ltd (Newport, UK), and tissue culture reagents were purchased from the Gibco Life Technologies (Dublin, Ireland).

D10

DMEM containing 4.5 g/L L-glucose, 10 % (v/v) heat inactivated foetal calf serum (FCS; Invitrogen), 250 units penicillin/streptomycin, 0.26 mg/mL L-Glutamine, and 97 mg/mL sodium pyruvate.

R10

RPMI containing 4.5 g/L L-glucose, 10 % (v/v) FCS, 250 units penicillin/streptomycin, 0.26 mg/mL L-Glutamine, and 97 mg/mL sodium pyruvate.

LB broth

20 g/L of LB low salt broth (Melford) in ddH_2O.

LB Agar

LB broth, 15 % (w/v) agar (Oxoid).

Ampicillin solution

100 mg/mL ampicillin sodium salt (Melford) in ddH_2O, sterilised through 0.22 μm filter.

FACS block

0.5 % BSA (w/v), 5 mM EDTA, 5 % rabbit serum (Merk), 4 ug/ mL 2.4G2 (BD Pharmingen, Catalog No. 553142) in PBS.

FACS buffer

2 % (v/v) FCS, 0.0 5% (w/v) sodium azide in PBS.

Saponin buffer

2 % (v/v) FCS, 0.05 % (w/v) sodium azide, 0.5 % (w/v) saponin (Acros Organics) in PBS.

Mammalian lysis buffer

100 mM TRIS pH 8.5, 200 mM NaCl, 5 mM EDTA, 0.2 % SDS.

2.2 Animal studies

2.2.1 Animals

All animal experiments were performed in accordance with the United Kingdom Home Office Animal (Scientific procedures) act of 1986. C57BL/6J male mice from Charles River (UK), were given free access to water and standard chow and housed in conventional cages (12 h light/dark cycle) in accordance with institutional guidelines at Cardiff University and 10-12-week-old C57BL/6 mice were sacrificed using a schedule one method (slow rising concentration of CO_2, and death was confirmed by cervical dislocation).

Mice with a myeloid deficiency of GATA6 (*Gata6*-KO^{mye} mice) and WT mice with the same genetic background were generated in Prof. Phil Taylor's lab, as previously described[69]. Briefly, conditional KO *Gata6*$^{tm2.1Sad/J}$ mice were crossed with LysozymeM (*Lyz2*) Cre-recombinase knockin mice to generate mice with a myeloid deficiency of GATA6[69].

2.2.2 Genotyping

A mouse ear biopsy was placed into a 1.5 mL Eppendorf tube containing 50 µL lysis buffer (100 mM Tris (pH 8.5), 5 mM ethylenediaminetetraacetic acid (EDTA), 0.2 % sodium dodecylsulphate (SDS) (w/v), 200 mM NaCl, 100 µg/ mL proteinase K – all from Sigma). The tube was then incubated overnight at 52 °C. The reaction was terminated by incubating the samples at 72 °C for 15 min. After cooling down, 400 µL of nuclease-free water was added to each tube to reduce the concentration of SDS, vortexed for 5 sec and left to settle.

These samples then underwent polymerase chain reaction (PCR) at 1 µL DNA per 20 µL reaction. The primers used are listed below.

Forward: CAGGCATGCTTTCTCTAGTCAG

Reverse: TGATCCTGGCAATTTCGGCT

GoTaq green polymerase (Promega, Southampton, UK) was used for conventional PCR reaction mixes, and the reaction mix was as follows: 1× Buffer

($MgCl_2$ final made up to 2.5 mM), 250 – 500 nM primers, 200 nM of each dNTP, 0.05 Units/µl GoTaq DNA polymerase. The final volume was 20 µl per reaction. The samples were inserted in a Mastercycler® nexus (Eppendorf Ltd, Stevenage, UK) and run according to the following program:

94°C - 3min

94°C - 20 sec

64°C - 30 sec * Reduce by -0.5°C every cycle

72°C - 35 sec

Repeat 2- 4 for 12 cycles

94°C - 20 sec

58°C – 30 sec

72°C – 35 sec

Repeat 6 – 8 for 25 cycles

72°C – 2 min

4°C ∞

2.2.3 Peritoneal lavage

Mice were killed by asphyxiation using CO_2, and death confirmed using cervical dislocation. Cadavers were placed on their back and sprayed with ethanol. The outer skin of the peritoneum was cut using scissors and forceps and gently pulled back to expose the inner skin lining the peritoneal cavity. 5 mL lavage solution (PBS + 5mM EDTA) was injected into the peritoneal cavity using a 21 G needle, being careful not to puncture any organs. After injection, the peritoneal cavity was massaged to dislodge any attached cells into the PBS solution, the lavage solution was slowly drawn back up into the syringe, avoiding the intestines and fat. The needle was then removed, lavage fluid was transferred to a 15 mL falcon tube and place on ice until further processing.

2.3 Isolation of pResMϕ

2.3.1 *Gata6*-WT and -KO^{mye} pResMϕ phenotyping and isolation

Peritoneal cells were recovered by lavage. Individual lavages were phenotyped using F4/80-APC (Bio-Rad, Kidlington, UK; MCA497APC) and CD73-ef450 (eBioscience, UK; 48-0731-80) and analysed by flow cytometry. For cell purifications, lavages from phenotyped mice were pooled and stained with F4/80-APC, CD73-ef450 (eBioscience, UK) and MHCII- PercpCy5.5 (BD biosciences, M5/114.15.2). Cells were sorted on a FACSAria III (BD biosciences, Berkshire, UK) for single events (via pulse width or FSC_{height} vs FSC_{area} respectively) with the defined flow-cytometric phenotype (Figure 2.1, Table 2.3). Cell purities were verified by flow cytometry.

a) Surface staining for phenotyping

Lavages (50 µL) were plated out onto a V-bottom 96 well plate (Fisher Scientific, Loughborough, UK). Cells were incubated at room temperature for 10 min with 50 µL anti-CD16/32 (BD biosciences, Berkshire, UK) diluted 1:500 in FACS block to block Fc receptors. This was followed by cell surface staining with 100 µL ABmix1 diluted in MACS buffer and incubated at 4 °C for 10 min. ABmix1 is described in Table 2.1.

b) Surface staining for isolation of Gata6-WT and -KO^{mye} pResMϕ by FACS

Pooled samples were transferred to a 15 mL falcon tube. Cells were then washed with FACS buffer and incubated at 4 °C for 40 min with anti-CD16/32 (BD biosciences, Berkshire, UK) diluted 1:500 in FACS block to block Fc receptors. This was followed by cell surface staining with ABmix2 diluted in MACS buffer and incubating at 4 °C for 40 minutes. ABmix2 antibodies are listed in Table 2.2.

Table 2.1 Antibodies used in flow cytometry for phenotyping of *Gata6*-WT and -KO^{mye} pResMϕ (ABmix1).

Antibody	Conjugate	Clone	Company	Dilution	Cat. number
F4/80	APC	Cl:A3-1	Serotec	1:10	MCA497APC
CD73	ef450	eBioTY/11.8	eBioscience	1:15	M5/114.15.2

Table 2.2 Antibodies used in flow cytometry for isolation of *Gata6*-WT and -KO^{mye} pResMϕ (ABmix2).

Antibody	Conjugate	Clone	Company	Dilution	Cat. number
F4/80	APC	Cl:A3-1	Serotec	1:200	MCA497APC
CD73	ef450	eBioTY/11.8	eBioscience	1:300	48-0731-80
MHCII	PercpCy5.5	M5/114	BD biosciences	1:400	M5/114.15.2

c) Phenotyping and purification of Gata6-WT and -KO^{mye} pResMϕ

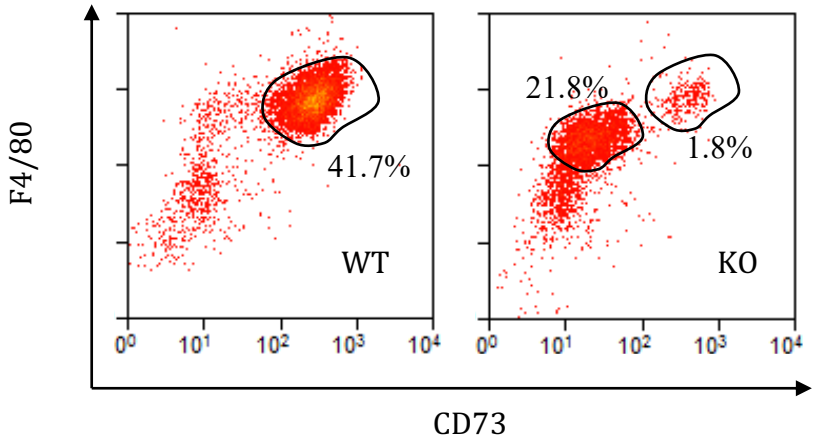


Figure 2.1 Flow-cytometric analysis of the pResϕ-*Gata6*-WT/KO^{mye} phenotyping.
Data are representative of one of two similar experiments with 1 female 11 week-old, C57BL/6 mice.

Table 2.3 Summary of *Gata6*-WT and -KO^{mye} pResMϕ used for LC-MS.

Cell Type	Abrev.	Purification criteria
Peritoneal tissue-resident macrophage wild type	pResMϕ-WT	F4/80^{high} MHCII^{low} CD73^{high}
Peritoneal tissue-resident macrophage Gata6-KO^{mye}	pResMϕ-Gata6-KO^{mye}	F4/80^{high} MHCII^{low} CD73^{low}

2.4 Cell culture

2.4.1 Culture of adherent cells

RAW 264.7 cells and human embryonic kidney (HEK293T) cells were cultured in D10 containing 1 % Penicillin-Streptomycin (GE Healthcare, Cardiff, UK)) at 37 °C and 5 % CO_2. Cells were thawed, added to 9 mL of D10 and centrifuged at 400 g for 5 min at 4 °C. The supernatant was discarded, and the cells resuspended in 1 mL of D10. The cells were added with 50 mL of D10 into a T175 cell culture flask.

Once a confluency was of 80 - 90 % was reached, the D10 was removed and the cells washed twice using PBS; 10 mL trypsin was added, and the cells incubated at 37 °C for 3 - 5 min. In order to stop the reaction, D10 was added and the cell suspension transferred into a 50 mL falcon tube and centrifuged at 400 g for 5 min at 4 °C. The supernatant was discarded, and the cells resuspended in 10 mL of D10. Cells were diluted into a new flask at required ratios and incubated at 37 °C and 5 % CO_2.

2.4.2 Culture of non-adherent cells

Jurkat cells were cultured in R10 containing 1 % Penicillin-Streptomycin (GE Healthcare, Cardiff, UK) at 37 °C and 5 % CO_2.

Cells were thawed, added to 9 mL of R10 and centrifuged at 400 g for 5 min at 4 °C. The supernatant was discarded, and the cells resuspended in 1 mL of R10. The cells were added to 50 mL of R10 into a T175 cell culture flask.

Once a confluency of 80 - 90 % was reached, the cell suspension was transferred into a 50 mL falcon tube and centrifuged at 400 g for 5 min at 4 °C. The supernatant was discarded, and the cells resuspended in 10 mL of R10. Cells were diluted into a new flask at required ratios and incubated at 37 °C and 5 % CO_2.

2.5 Molecular biology

2.5.1 Agarose gel electrophoresis (AGE)

Agarose was dissolved in 0.5 x Tris-EDTA (TBE) solution (1 – 2 % (w/v) agarose) by heating the solution in a microwave until boiling. Once the gel had cooled sufficiently, ethidium bromide (100 ng/mL) was mixed into the solution, which was poured into a gel caster with well moulds. Once the gel was set, it was placed in a gel tank filled with 0.5 x TBE, and the samples (with DNA loading dye (Promega)) and DNA ladders (100 bp (base pair) (Promega)) were loaded into the wells. The gel was run at 100 volts for 30 minutes and viewed under UV light source.

2.5.2 In-fusion cloning

a) Primer design

Primers used for in-fusion cloning were designed, using Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/), to include a 15 base pair (bp) region which overlapped with the regions flanking the restriction endonuclease digestion site in the vector sequence, and to have a melting temperature 55 – 65 °C. All primers and oligonucleotides were obtained from Sigma as salt free purified lyophilised DNA and stored at -20 °C (Table 2.4).

b) Vector linearization

SU6EW vector (3 µg) was mixed with the restriction enzyme PmeI (5 µL) and 10 x CutSmart NEB buffer (5 µL), and the total volume was made up to 50 µl with RNAse-free water. The tubes were placed in a PTC-200 thermal cycler and incubated at 37 °C for 1 hour followed by a 20 min step at 65 °C. PCR products were then run on a 1 % Agarose gel with a 1 kb ladder (Promega). DNA from the tubes with linearized vector was then purified using a Macherey-Nagel kit and quantified using a Nanodrop spectofotometer (Thermo Scientific). Purified linearized vector was stored at -20 °C.

Table 2.4 In-fusion cloning primers.

Table showing in-fusion cloning primers and the vectors/ restriction endonucleases used to produce new vectors.

Starting plasmid	Restriction	New Plasmid	Primer Pairs
SU6EW	Pme I	CTSE shRNA1	Forward:
			Agaaaagccttgtttg**CGAGTCCTGTAATGTGTATTCTCGAGAATACACATTACAGGACTCGC**TTTTTGTTTaaacGAATTCctgca
			Reverse:
			TGCAGGAATTCGTTTAAACAAAAAGCGAGTCCTGTAATGTGTATTCTCGAGAATACACATTACAGGACTCGCAAACAAGGCTTTTCT
		CTSE shRNA2	Forward:
			Agaaaagccttgtttg**CAGAACTTCACTGTCATCTTCTCGAGAAGATGACAGTGAAGTTCTGC**TTTTTGTTTaaacGAATTCctgca
			Reverse:
			TGCAGGAATTCGTTTAAACAAAAAGCAGAACTTCACTGTCATCTTCTCGAGAAGATGACAGTGAAGTTCTGCAAACAAGGCTTTTCT
		CTSE shRNA3	Forward:
			Agaaaagccttgtttg**GCCCGACAAGATCAAACATCTTCTCGAGAAGATGTTTGATCTTGTCGGGC**TTTTTGTTTaaacGAATTCctgca
			Reverse:
			TGCAGGAATTCGTTTAAACAAAAAGCCCGACAAGATCAAACATCTTCTCGAGAAGATGTTTGATCTTGTCGGGCAAACAAGGCTTTTCT
		GBA2 shRNA1	Forward:
			Agaaaagccttgtttg**TACAAATCTGCACTGTTCAACTCGAGTTGAACAGTGCAGATTTGTACTTTTTGTTT**aaacGAATTCctgca
			Reverse:
			TGCAGGAATTCGTTTAAACAAAAAGTACAAATCTGCACTGTTCAACTCGAGTTGAACAGTGCAGATTTGTACAAACAAGGCTTTTCT
		GBA2 shRNA2	Forward:
			Agaaaagccttgtttg**AAGTATTTCCTACCCTGCATCTCGAGATGCAGGGTAGGAAATACTTCTTTTTGTTT**aaacGAATTCctgca
			Reverse:
			TGCAGGAATTCGTTTAAACAAAAAGAAGTATTTCCTACCCTGCATCTCGAGATGCAGGGTAGGAAATACTTCAAACAAGGCTTTTCT
		GBA2 shRNA3	Forward:
			Agaaaagccttgtttg**CCGAAAGAGATTCCCAAGTTCTCGAGAACTTGGGAATCTCTTTCGGCTTTTTGTTT**aaacGAATTCctgca
			Reverse:
			TGCAGGAATTCGTTTAAACAAAAAGCCGAAAGAGATTCCCAAGTTCTCGAGAACTTGGGAATCTCTTTCGGCAAACAAGGCTTTTCT
		SMPD1 shRNA1	Forward:
			Agaaaagccttgtttg**GTGTACTGGACGGGAGATATCTCGAGATATCTCCCGTCCAGTACACCTTTTTGTTT**aaacGAATTCctgca
			Reverse:
			TGCAGGAATTCGTTTAAACAAAAAGGTGTACTGGACGGGAGATATCTCGAGATATCTCCCGTCCAGTACACCAAACAAGGCTTTTCT
		SMPD1 shRNA2	Forward:
			Agaaaagccttgtttg**GAGCCTCCCAGATGCTAATACTCGAGTATTAGCATCTGGGAGGCTCCTTTTTGTTT**aaacGAATTCctgca
			Reverse:
			TGCAGGAATTCGTTTAAACAAAAAGGAGCCTCCCAGATGCTAATACTCGAGTATTAGCATCTGGGAGGCTCCAAACAAGGCTTTTCT
		SMPD1 shRNA3	Forward:
			Agaaaagccttgtttg**CCGAAAGAGATTCCCAAGTTCTCGAGAACTTGGGAATCTCTTTCGGCTTTTTGTTT**aaacGAATTCctgca
			Reverse:
			TGCAGGAATTCGTTTAAACAAAAAGCCGAAAGAGATTCCCAAGTTCTCGAGAACTTGGGAATCTCTTTCGGCAAACAAGGCTTTTCT

c) Preparation for cloning

Linearised vector (200 ng) was mixed with insert (100 ng) and InFusion cloning reaction enzyme (2 µL) (Clontech), total volume was made up to 10 µL with water. The tubes were placed in a thermocycler (Mastercycler® nexus, Eppendorf) and incubated at 50 °C for 15 minutes for the recombination reaction.

d) E. coli transformation via heat shock

Diluted In-fusion reaction solution (1 µL) (described in 2.5.2c) was added to a 50 µL aliquot of one-shot Top10 competent E. coli cells and incubated on ice for 30 minutes. The tubes were placed into a ready PTC-200 thermal cycler set at 42 °C and held for exactly 30 sec before putting them back on ice for 5 min. Pre-warmed SOC media (950 µL) was added to the transformed cells before incubating in a shaking heating block for 1 h at 37 °C at 250 rpm. Transformed bacteria were then plated on LB-Amp (pre-warmed to 37 °C) plates and incubated overnight at 37 °C.

e) Colony PCR

Colony PCR mix was prepared by adding 8.4 µL Kappa enzyme, 0.125 µL each primer (forward and reverse), and 16.5 µL H_2O. Selected colonies (14 per plasmid) were added to the mix and PCR products were synthesised using a Mastercycler® nexus thermocycler (Eppendorf) under the following parameters:

1. Incubate at 94 °C for 2 minutes
2. Incubate at 94 °C for 15 seconds
3. Incubate at 55 °C for 30 seconds
4. Incubate at 72 °C for 3 minutes
5. Go to step 2 for 34 cycles
6. Incubate at 4 °C and pause

The PCR product was then run on a 1.5 % agarose gel alongside a sample of 100 bp DNA ladder (Promega) and visualised using a Fluoro-S Multimager (Bio- rad).

f) E. coli culture

LB broth was prepared by adding LB tablets to water in a glass bottle and autoclaving it. Selected positive *E.coli* colonies were touched with the end of a pipette tip and placed into a 50 mL Falcon tube with 10 mL LB broth with ampicillin (100 μg/mL). The culture was then incubated at 37 °C in a shaking incubator (225 rpm) for 18 h.

g) Plasmid DNA extraction

Qiagen Mini and Maxi plasmid prep kits were used to extract plasmid DNA from *E.coli*. Bacterial cultures were pelleted by centrifugation at 6,800 x g for 3 min at room temperature. The cell pellets were then resuspended in lysis buffer to release plasmid DNA, and lysis was stopped after 5 minutes to minimize contamination with bacterial DNA. The resulting solution was centrifuged for 10 min at 17,900 x g. The supernatant was placed onto a membrane (provided in the Kit) to bind the DNA, which was then washed with ethanol. The purified plasmid DNA was eluted with elution buffer. DNA concentrations were measured using a Nanodrop and DNA stored at -20 °C.

h) DNA sequencing

Plasmid DNA was sent for sequencing by Eurofins Genomics, and results were manually checked. Selected plasmids were then used for lentivirus production.

2.5.3 Lentivirus production

a) Transfection

HEK293T cells (10 x10^6 cells) were seeded, in a T175 flask at 37 °C, to reach 70 – 90 % confluency after 16 – 20 h. Cells were washed, taking care not to disturb the monolayer. And then 12 mL of fresh complete media added. The flask was placed back into the incubator for 5 min (37 °C).

Cells were transfected using the Effectene transfection kit (Qiagen). Briefly, 2 µg of lentiviral plasmid was mixed with the pCMVΔ8.91 (1.5 µg) and pMD2G (1 µg) plasmids, and this solution was made up to 600 µL with condensation buffer EC. The Enhancer (36 µL) reagent provided in the kit was added to the mix, vortexed and incubated at RT for 5 minutes to condense the DNA. Effectene transfection reagent (120 µL) was added, vortexed and incubated at RT for 10 minutes, to form micelle complexes around the condensed DNA. 2.6 mL D10 media was then added to the solution, gently mixed and poured dropwise into the cell flask directly to the monolayer. Cells were incubated for 48 h at 37 °C. Cell supernatant was collected and transferred to a 50 mL falcon tube per flask.

b) Viral particle concentration by sucrose gradient

The supernatant collected from the transfection step was filtered through a 0.45 µm sterile millex®GP filter (Millipore Ireland Ltd.) into a sterile 50 mL tube. Filtered medium was then carefully overlaid on top of 3 mL of 20 % sucrose (Sigma) gradient in Beckman ultracentrifuge conical tubes. And centrifuged at 90,000 g for 90 min at 4 °. The pellet was resuspended in 1 mL AIM-V media and aliquoted, then stored at -80 °C.

2.5.4 Lentivirus testing

Jurkat cells (0.2 x 10^6 cells per well) were seeded into a 24-well plate with a titrated volume (0 – 10 µL) of concentrated virus with R10 kept constant (200 µL). The plates were incubated at 37 °C in a humidified incubator. Cells were topped up with 500 µL of media on day 1 then harvested on day 3, washed three times and fixed for 5 min with 2 % formaldehyde (Sigma). Cells were then

analysed using flow cytometry and the percentage of infected (GFP^+) cells calculated.

2.5.5 Use of lentiviral particles in culture

RAW 264.7 cells were plated into a 12-well plate (0.4×10^6 cells in 400 µL D10) with a titrated volume (0 – 100 µL) of virus and the plate incubated for 5 days at 37 °C. Media (500 µL) was added to wells every 24 hours.

2.5.6 Real-time PCR

a) Primer design

Primers used for real-time PCR were designed to be intron-spanning, and to have a melting temperature 55 - 65 °C. The intron-spanning aspect ensured that only mRNA and not genomic DNA were amplified. All primers and oligonucleotides were obtained from Sigma as salt free purified lyophilised DNA. The sequences of the real-time PCR primers used in this thesis were:

β-actin	F: GACCTCTATGCCAACACAGT	R: AGTACTTGCGCTCAGGAGGA
Gata6	F: AAAGCTTGCTCCGGTAACAG	R: TCTCCCACTGCAGACATCAC
Ywhaz	F: TGCAAAAACAGCTTTCGATG	R: CCTGCTTCTGCTTCATCTCC
Ctse	F: GGGTCTGGGATACCCCTCAT	R: GTCTCCCACCTGGATTCCATC
Smpd1	F: AGAGCACTCCTGTCAATGGC	R: AAGCCCCCAATTCTTAGGGTG
Gba2	F: TCATGACTACCAGGACAGCAG	R: CGGAAGGGCTCATTCCACAA

b) RNA extraction

RNA was extracted using the RNeasy mini kit (Qiagen). Briefly, lysis buffer was added to cell pellets to free nucleic acids, which were brought into solution with ethanol and bound to a membrane. The nucleic acids were washed to remove chromatin contaminants, and contaminating salts removed using ethanol. Last, the ethanol was removed, and RNA eluted with RNAse-free dH_2O (Life Technologies).

RNA concentration was measured using the Nanodrop spectrophotometer (Thermo Scientific). RNA concentration was determined based on the equation: optical density at 260 nm × 40 ng/µL (constant) = ng/µL of RNA.

c) Reverse transcription

Reverse transcription of RNA into complementary DNA (cDNA) was carried-out using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). Briefly, RNA obtained above (1000 ng per reaction) serves as the template in the reverse transcription reaction, where random hexamers were used to synthesise cDNA.

d) Real-time PCR (qPCR)

Power SYBR® Green PCR Master Mix (Thermo Fisher scientific) was used for real-time PCR. Power SYBR Green PCR Master Mix (10 µL) was added to dH_2O (7.5 µL), and 1 µl primer mix (0.5 µg/mL each primer) to make a gene specific master mix; cDNA (1.5 µL) from each sample was added to the reaction mix (17.5 µL) in a 96-well plate well. The plate was inserted into a QuantStudio 3 Real-Time PCR system and the PCR amplification started as per instrument instruction:

50°C - 2min
95°C - 10 min
95°C - 15 sec
60°C - 1 min
Repeat 3- 4 for 40 cycles
95°C - 15 sec
4°C ∞

Ct values were calculated for every gene specific reaction and normalised against the housekeeping gene (*Ywhaz* and *ß-actin*). These values were expressed as a fold change (using the ΔΔ-Ct method[155]).

2.6 Mass Spectrometry

2.6.1 Chemicals

Lipid standards (Table 2.5) were purchased from Cayman Chemical (Cambridge, UK). HPLC grade solvents and acids were purchased from Thermo Fisher Scientific Ltd (Newport, UK). All other reagents were obtained from Sigma-Aldrich unless otherwise stated.

2.6.2 Lipid Extraction

a) Global lipidomics

Arachidonic acid d_8 and PE ether (Table 2.5) were added to each lavage sample before extraction (2.5 ng), as internal standards. Lipids were extracted by adding a solvent mixture (1 M acetic acid/2-isopropanol/hexane (2:20:30, v/v) to the sample at a ratio of 2.5 mL of solvent mixture/1 mL of sample, vortexing for 1 min, and then adding 2.5 mL hexane. Following vortexing and centrifugation (250 g for 5 min), lipids were recovered in the upper hexane layer. The samples were then re-extracted by addition of an equal volume of hexane, followed by vortexing and centrifugation (250 g for 5 min). The bottom layer was again re-extracted, using the Bligh and Dyer extraction method as follows: 3.75 mL chloroform:MeOH (1:2, v/v) was added, sample was vortexed for 1min, and then 1.25 mL chloroform added. Samples were vortexed for 30 seconds then 1.25 mL water. Following vortexing and centrifugation (250 g for 5 min), lipids were recovered in the organic phase. The combined chloroform and hexane phases containing lipids were dried in a vacuum drier, then redissolved in 200 µL methanol and stored at -80 °C until analysis.

Table 2.5 Lipid standards.

Standard	Formal Name	Exact Mass
PGE_2-d_4	prostaglandin E_2-d_4	356.2501
TxB_2-d_4	Thromboxane B_2-d_4	374.2606
11-dehydro-TxB_2-d_4	11-dehydro-Thromboxane B_2-d_4	372.245
13-HODE-d_4	13S- hydroxy- 9Z, 11E- octadecadienoic- 9, 10, 12, 13- d_4 acid	300.2603
12-HETE-d_8	12S- hydroxy- 5Z, 8Z, 10E, 14Z-eicosatetraenoic- 5, 6, 8, 9, 11, 12, 14, 15-d_8 acid	328.2854
15-HETE-d_8	15S- hydroxy- 5Z, 8Z, 11Z, 13E-eicosatetraenoic- 5, 6, 8, 9, 11, 12, 14, 15-d_8 acid	328.2854
Arachidonic acid-d_8	5Z,8Z,11Z,14Z-eicosatetraenoic-5,6,8,9,11,12,14,15-d_8 acid	312.29044
PE Ether (PEE (40:1))	1,2-di-(3,7,11,15-tetramethylhexadecyl)-sn-glycero-3-phosphoethanolamine	775.6819

b) Sphingolipidomics

Lipid standards (Cer/Sph Mixture II, Avanti polar lipids, Table 2.6) were added to each lavage sample before extraction, as internal standards. Lipids were extracted by adding a solvent mixture (1 M acetic acid/2-isopropanol/hexane (2:20:30, v/v) to the sample at a ratio of 2.5 mL solvent mixture/1 mL of sample, vortexing, and then adding 2.5 mL hexane. Following vortexing and centrifugation, lipids were recovered in the upper hexane layer. The samples were then re-extracted by addition of an equal volume of hexane, followed by vortexing and centrifugation. The bottom layer was again re-extracted, using the Bligh and Dyer method, adding 3.75 mL of chloroform:MeOH (1:2, v/v), vortexing for 1 min, and added 1.25 mL chloroform, vortexing for 30 seconds and added 1.25 mL water[156]. Following vortexing and centrifugation, lipids were recovered in the organic phase. The combined chloroform and hexane phases containing lipids were dried in a vacuum drier redissolved in 200 µL methanol and stored at -80 °C until analysis.

2.6.3 Targeted analysis of lipids

a) Fatty acids and eicosanoids

Lipids extracts were separated by reverse phase HPLC using a C18 Spherisorb ODS2, 5 µm, 150 x 4.6 mm column (Waters, Hertfordshire, UK). With a mobile phase gradient of 50 – 90 % solvent B (60 % methanol, 40 % acetonitrile, 0.1 % glacial acetic acid) in solvent A (75 % water, 25 % acetonitrile, 0.1 % glacial acetic acid) over 20 min, 90 % solvent B over 5 min, 90 – 50 % solvent B from 25 – 25.1 min and 50 % solvent B from 25.1 – 30 min with a flow rate of 1 mL/min. LC-MS/MS was performed using a AB Sciex 6500 Q-Trap, monitoring the parent to daughter *m/z* as listed in Table 2.7. Peak areas were then used to calculate extraction efficiencies.

Table 2.6 SL standards (Cer/Sph Mixture II, Avanti polar lipids).

Lipid	Q1	Q3
C17 So	286.45	268.2
C17 Sa	288.44	270.2
C17 So-1-P	366.34	250.2
C17 Sa-1-P	368.41	270.29
C12 Cer	482.59	264.2
C25 Cer	664.83	646.7
C12 Cer-1-P	562.6	264.5
C12 SM	647.67	184.1
C12 GlucCer	644.68	264.5
C12 LacCer	806.69	264.4

Table 2.7 Parent to daughter ion transitions used for measuring fatty acids and eicosanoid in murine samples.

Q1 Mass (Da)	Q3 Mass (Da)	Lipid
355.2	275.1	PGE_2-d_4
373.2	173.1	TxB_2 d_4
371.2	165.1	11 dehydro TxB_2 d_4
299.2	198.1	13-HODE d_4
327.2	184.1	12-HETE d_8
327.2	226.1	15-HETE d_8
327.2	116.1	5-HETE-d_8

b) Sphingolipidomics

i) Cer, SM, monohexosylceramides (HexCer) and dihexosylceramides (LacCer)

Lipids extracts were analysed using a reverse phase Accucore C18 column (150 x 2.1 mm, 2.6 µm) using an isocratic mobile phase A (H_2O:ACN 80:20 v/v) and a phase B (IPA:ACN, 70:30 v/v) at a flow rate of 0.4 mL/min over 65 min, with 4 mM ammonium acetate. Samples were maintained at 4 °C and the column at 25 °C. LC- MS/MS was performed with a HPLC (Shimadzu, Japan) and 6500 Q-Trap (AB Sciex, Canada), using MRMs in positive ionisation mode to monitor the lipids listed in Table 2.8, in Table 2.9, and Table 2.10.

ii) Resolution of GlcCer and galactosylceramide (GalCer)

These compounds were analysed by normal phase LC, using a Supelco 2.1 (i.d.) × 250 mm LC-Si) and an isocratic elution with mobile phase A (CH_3CN/CH_3OH/HCOOH, 97/2/1, v/v/v, with 5 mM ammonium formate) at a flow rate of 1.5 mL/min over 8 min. LC/MS/MS was performed with a HPLC (Shimadzu, Japan) and 6500 Q-Trap (AB Sciex, Canada), using MRMs in positive ionisation mode to monitor the lipids listed Table 2.9.

2.6.4 Untargeted/Global lipid analysis

Lipid extracts were separated on an Accucore C18 column (150 x 2.1 mm, 2.6 µm) using solvent mixture of 84 % mobile phase A (H_2O:ACN 80:20 v/v) and 16 % phase B (IPA: ACN, 70:30 v/v) with 4 mM ammonium acetate. The A/B ratio was maintained at 84/16 for 12 min, followed by a linear gradient of 60 % B for 7 min, 72 % B for 23 min, 84 % B for 9 min, and 100 % B which was held for 6 min, followed by an 8 min wash of the column at 84/60 A/B before the next run. Samples were maintained at 4 °C and the column maintained at 25 °C. High resolution (60,000, FWHM at 400 amu) full-scan spectra were acquired in both +ve and –ve ESI mode over 100 – 1800 *m/z* range on an Orbitrap (Elite) MS. The MS conditions were as follows: HESI-II temperature 350 °C, N_2 as drying gas, sheath gas flow 52 arbitrary units, auxiliary gas flow 18 units, capillary temp 320 °C, spray voltage +/- 3.5 kV and S-lens RF level 65.60 % respectively for

negative and positive ion mode. A one-off column conditioning at the beginning was done with two runs of solvent blank followed by runs of solvent blank using same mobile phase gradient.

Table 2.8 Parent to daughter ion transitions used for measuring Cer species.

Lipid	Q1 Mass (Da)	Q3 Mass (Da)
C17:0 Sa	288.44	270.2
C17:0 Sa-1-P	368.4	252.2
C17:1 So	286.45	268.2
C17:1 So-1-P	366.3	250.2
C18:0 Sa	302.4	252.2
C18:0 Sa-1-P	382.4	266.4
C18:1 So	300.4	282.4
C18:1 So-1-P	380.4	264.4
Cer 18:1/12:0	482.6	264.2
Cer 18:1/16:0	538.7	264.4
Cer 18:1/18:0	566.5	264.4
Cer 18:1/18:1	564.5	264.4
Cer 18:1/20:0	594.6	264.4
Cer 18:1/20:1	592.6	264.4
Cer 18:1/22:0	622.6	264.4
Cer 18:1/22:1	620.6	264.4
Cer 18:1/24:0	650.6	264.4
Cer 18:1/24:1	648.6	264.4
Cer 18:1/25:0	664.8	264.4
Cer 18:1/26:0	678.7	264.4
Cer 18:1/26:1	676.7	264.4
CerP 18:0/24:0	732.9	266.4
CerP 18:0/24:1	650.6	266.4
CerP 18:1/12:0	562.5	264.5
CerP 18:1/16:0	618.7	264.4
CerP 18:1/24:0	730.9	264.4

Table 2.9 Parent to daughter ion transitions used for measuring monohexosylceramides.

Lipid	Q1 Mass (Da)	Q3 Mass (Da)
HexCer 18:0/14:0	674.5	266.4
HexCer 18:0/14:1	672.5	266.4
HexCer 18:0/16:0	702.6	266.4
HexCer 18:0/16:1	700.6	266.4
HexCer 18:0/18:0	730.6	266.4
HexCer 18:0/18:1	728.6	266.4
HexCer 18:0/20:0	758.6	266.4
HexCer 18:0/20:1	756.6	266.4
HexCer 18:0/22:0	786.7	266.4
HexCer 18:0/22:1	784.6	266.4
HexCer 18:0/24:0	814.9	266.4
HexCer 18:0/24:1	812.7	266.4
HexCer 18:0/26:0	842.7	266.4
HexCer 18:0/26:1	840.7	266.4
HexCer 18:1/12:0	644.7	264.5
HexCer 18:1/14:0	672.5	264.4
HexCer 18:1/16:0	700.6	264.4
Hexcer 18:1/16:1	698.5	264.4
Hexcer 18:1/18:0	728.6	264.4
Hexcer 18:1/18:1	726.6	264.4
HexCer 18:1/20:0	756.6	264.4
Hexcer 18:1/20:1	754.6	264.4
HexCer 18:1/22:0	784.6	264.4
Hexcer 18:1/22:1	782.6	264.4
HexCer 18:1/24:0	812.7	264.4
Hexcer 18:1/24:1	810.7	264.4
Hexcer 18:1/26:0	840.7	264.4
Hexcer 18:1/26:1	838.7	264.4

Table 2.10 Parent to daughter ion transitions used for measuring LacCer and SM.

Lipid	Q1 Mass (Da)	Q3 Mass (Da)
LacCer 18:1/12:0	806.7	264.4
LacCer 18:1/16:0	862.7	264.4
LacCer 18:1/18:0	890.7	264.4
LacCer 18:1/18:1	888.6	264.4
LacCer 18:1/20:0	918.7	264.4
LacCer 18:1/20:1	816.7	264.4
LacCer 18:1/22:0	946.8	264.4
LacCer 18:1/22:1	844.7	264.4
LacCer 18:1/24:0	974.8	264.4
LacCer 18:1/24:1	972.7	264.4
LacCer 18:1/26:0	1002.8	264.4
LacCer 18:1/26:1	1000.8	264.4
SM 18:0/12:0	649.7	184.1
SM 18:0/14:0	677.6	184.4
SM 18:0/16:0	705.6	184.4
SM 18:0/18:0	733.8	184.4
SM 18:0/20:0	761.6	184.4
SM 18:0/22:0	789.7	184.4
SM 18:0/24:0	817.9	184.4
SM 18:0/26:0	845.7	184.4
SM 18:1/12:0	647.7	184.1
SM 18:1/14:0	675.5	184.4
SM 18:1/16:0	703.6	184.4
SM 18:1/16:1	701.6	184.4
SM 18:1/18:0	731.8	184.4
SM 18:1/18:1	729.6	184.4
SM 18:1/20:0	759.6	184.4
SM 18:1/20:1	757.6	184.4
SM 18:1/22:0	787.6	184.4
SM 18:1/22:1	785.6	184.4
SM 18:1/24:0	815.9	184.4
SM 18:1/24:1	813.7	184.4
SM 18:1/26:0	843.7	184.4
SM 18:1/26:1	841.7	184.4

2.6.5 Data normalisation by DNA content

a) DNA purification

Genomic DNA from pResMϕ was isolated using the PureLink Genomic DNA Mini Kit (Thermofisher) according to supplier's specifications. Briefly, cell lysates from pResMϕ resuspended in PBS were prepared by adding 20 µL proteinase K and 20 µL RNase A to the sample and incubated at room temperature for 2 min; added 200 µL Purelink Genomic Lysis (binding buffer) and incubated at 55 °C for 10 minutes, followed by the addition of 200 µL of 100 % ethanol to the lysate. The sample was then applied to a PureLink spin column and washed twice with wash buffer 1 and 2, before DNA was eluted using elution buffer.

b) DNA quantification

Genomic DNA from pResMϕ was quantified using the Quant-iT dsDNA Assay Kit (ThermoFisher) according to supplier's specifications. Briefly, assay components were equilibrated at room temperature, and working solution was prepared by diluting the Quant-iT dsDNA reagent 1:200 in Quant-iT dsDNA buffer. Working solution was loaded (200 µL) in each well, 10 µL of each of the Quant-iT dsDNA standards was added, and unknown DNA sample to separate wells in duplicate. Fluorescence was measured using a microplate reader (excitation/emission maxima ~502/503 nm) and standard curve was used to determine DNA concentration.

2.7 Data analysis

2.7.1 Pathway analysis

The microarray data was previously published by Prof. Phil Taylor's lab[69]. The networks and functional analyses were generated through the use of QIAGEN's Ingenuity Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity).

2.7.2 XCMS processing

The .RAW files produced from the LC-MS runs were converted to .mzML format applying ProteoWizard 2.1[157] followed by deconvolution and peak alignment applying XCMS using a previously described method[158]; Results were then queried against two databases, Human Metabolome Database (HMDB) and LIPID MAPS using the database search program MS search with mass tolerance of 5 ppm. MS search enables the query of a list of *m/z* values against MS databases. Thus, providing the user with putative identifications for the features whose *m/z* values matched entries on the database.

2.7.3 Statistical analysis

XCMS-analysed data were processed using both an univariate and multivariate approach. Significantly different lipids were identified using an unpaired parametric t-test (Student t-test, unequal variances). P-values were adjusted for multiple comparisons, using the Benjamini-Hochberg method. Differential expression was classified as significant when $P < 0.05$ and $|FC| > 1.5$. Unsupervised PCA and supervised OPLS methods were used to obtain group clusters using the program Simca (version 14.0.1.0, Umetrics AB, Umeå, Sweden, 2013).

Statistical analysis of differential lipid levels, from a targeted analysis, consisted of a univariate model to detect individual lipids known to be present in Mϕ that are significantly different in abundance in two sample sets. P-values were adjusted for multiple comparisons, using the Benjamini-Hochberg method. P-values of less than 0.05 were considered statistically significant. Each row was analysed individually, without assuming a consistent SD, as indicated on legends. * $p< 0.05$ was considered statistically significant and ** indicated $p< 0.01$.

Chapter 3 GATA6 transcription factor regulates the expression of genes related to lipid metabolism

3.1 Introduction

3.1.1 *Gata6* in peritoneal Mϕ

As described in Chapter 1, the functions and phenotypes of ResMϕ are highly heterogeneous and, in most cases, the signals and molecular mechanisms that underlie their heterogeneity are poorly understood. In particular, the transcription factor GATA6 has a central role in promoting tissue-dependent characteristics of pResMϕ (*e.g.* phenotype, localization, proliferative renewal and function[159]), has previously been shown to be selectively expressed in these cells[51,69], and to regulate the expression of 'pResMϕ-specific genes' (PMSGs)[159]. Importantly, the transcriptional analysis of *Gata6*-KO^{mye} and -WT pResMϕ showed several differentially expressed genes involved in lipid metabolism[69]. Among the first genes identified were *Abca*, *Alox15*, *Sptlc2*, *Fabp4* and *Ldlr*. Such changes in gene expression are presumably linked to alterations in lipid metabolism related pathways.

In this chapter I will conduct a detailed analysis of the microarray data of Mϕ from *Gata6*-KO^{mye} and -WT mice to identify the gene networks regulated by GATA6 associated with lipid metabolic processes[69]. This is crucial in addressing the question of the relevance of GATA6 in regulating lipid biology in pResMϕ.

I will begin by assessing the changes in gene expression that occur in GATA6-deficient pResMϕ, which will result in a long list of genes with significantly altered transcript levels. Then, I will characterise the mechanisms that underlie the observed changes, in order to understand their biological relevance[160]. To achieve this, I will use a combination of gene-ontology rankings, pathway mappings and network analysis.

3.1.2 Aims

The studies described in this chapter will:

- Use ingenuity pathway analysis (IPA) to analyse gene pathways and networks regulated by GATA6, that are associated with lipid metabolic processes.

3.2 Results

Microarray data generated as part of a previous study by the Taylor lab, comparing pResMϕ from *Gata6*-KO^{mye} and -WT naïve mice was used to identify genes regulated by GATA6[69]. The microarray data was analysed using ingenuity pathway analysis (IPA) software. IPA provides two powerful tools to help determine biological mechanisms behind the changes in transcript levels, termed networks and canonical pathways, which will described in more detail in the following sections.

The analysis of this high-throughput gene-expression data was greatly facilitated by the availability of prior biological knowledge. Typically, such analysis is done using statistical gene-set-enrichment methods where differentially-expressed genes are interrogated using lists of genes associated with a particular biological function or pathway.

3.2.1 Canonical Pathways

The raw microarray data files were processed through an automated feature extraction and normalization pipeline developed at Cardiff University by Dr. Robert Andrews, and significantly-differentially expressed genes (p-value<0.05 after Benjamini-Hochberg correction and fold change≥ 1.5) were uploaded and analysed in IPA. The data analysis was performed in two primary steps, identification of significantly altered (i) canonical pathways, and (ii) functions/mechanisms related to lipid metabolism.

The first step of my analysis consisted of identifying which canonical pathways were significantly changed. The significance of the association between the dataset and the canonical pathway was established based on two considerations: (i) the first is the ratio between the number of genes from the dataset that match the genes in the pathway, divided by the total number of genes comprising the pathway; (ii) secondly a P-value, calculated using Fisher's exact test, determining the likelihood that the association between genes in the dataset and the canonical pathway is a random occurrence[160].

Using this analysis, I found that overall, the majority of dysregulated pathways were related to metabolic processes and a high proportion of pathways related to lipid metabolism were noted to be significantly enriched (Table 3.1.), either by affecting lipid pathways directly, or by altering lipid catabolism and transport, both of which are essential in maintaining lipid homeostasis (*i.e.* lysosomal metabolism). The most significant pathways are compared in Table 3.1. The 'Macropinocytosis Signalling', 'Phospholipase C Signalling', and 'Sphingosine-1-phosphate Signalling' pathways are the only ones for which IPA was able to calculate z-scores, and they are all predicted to be activated (positive z-scores). Thus, going forward, the goal of this chapter is two-fold, to assess which lipid classes were the most affected by GATA6 deficiency, and to gauge if lysosomal metabolism related genes could potentially contribute to alterations in lipid metabolism. In the next section the pathways identified above will be divided into 2 groups, depending on their function and expanded to categorise the biological/biochemical effect of these pathways. The first group will be based on the regulation of different lipid categories, whilst the second comprises the pathways involved in lysosome fusion and function. The data presented in the next 2 sections is merely descriptive of the changes observed, but following that I will continue with a more targeted analysis, focusing on SL and lysosomal metabolism.

Table 3.1 - The most significantly enriched lipid metabolism (ranked by log p-value) related pathways in *Gata6*-KO^{mye} pResMϕ according to IPA[1].

Ingenuity Canonical Pathways	-log(p-value)	Ratio	z-score
Macropinocytosis Signalling	5.877	0.25	0.632
Clathrin-mediated Endocytosis Signalling	4.94	0.157	NaN
Phospholipase C Signalling	4.06	0.135	0.392
Caveolar-mediated Endocytosis Signalling	3.18	0.183	NaN
Phagosome formation	3.13	0.156	NaN
Phagosome maturation	2.66	0.142	NaN
Sphingosine-1-phosphate Signalling	2.29	0.138	0.258
Prostanoid Biosynthesis	1.75	0.333	NaN
Superpathway of D-myo-inositol (1,4,5)-trisphosphate Metabolism	1.74	0.208	NaN

[1] A positive z-score indicates a predicted activation, and a negative z-score indicates a predicted inactivation of the enriched pathway. NaN, a z-score cannot be calculated for all IPA canonical pathways.

a) Regulation of specific lipid categories:

The canonical pathways identified by the IPA data analysis as significantly changed in *Gata6*-KO pResMϕ (Table 3.1.), were grouped within specific lipid categories involved in lipid metabolism and are listed below.

i) Phospholipid metabolism

Phospholipase C (PLC) Signalling - PLC selectively catalyses the hydrolysis of the phospholipid phosphatidylinositol 4,5-bisphosphate (PIP2) on the glycerol side of the phosphodiester bond. This reaction leads to the formation of inositol 1,2-cyclic phosphodiester, and the release of diacyl glycerol (DAG). The inositol intermediate is then hydrolyzed to inositol 1,4,5-trisphosphate (IP3)[92].

Superpathway of D-myo-inositol (1,4,5)-trisphosphate Metabolism - D-myo-inositol (1,4,5)-trisphosphate is a secondary messenger molecule used in signal transduction and lipid signaling. It is synthesized by hydrolysis of phosphatidylinositol 4,5-bisphosphate (PIP2), a phospholipid that is located in the plasma membrane, by the action of phospholipase C[161].

ii) Fatty acid metabolism

Prostanoid Biosynthesis - Prostanoids are cyclic, oxygenated products of ω3 and ω6 20- and 22-carbon essential fatty acids (FAs) that are formed enzymatically through cyclooxygenases[162].

iii) SL metabolism

Sphingosine-1-phosphate (S1P) Signalling (Figure 3.1) – S1P is a critical regulator of many physiological and pathophysiological processes, including cancer, atherosclerosis, diabetes and osteoporosis; it has also been shown that S1P has important intracellular targets involved in inflammation, cancer and Alzheimer's disease[162].

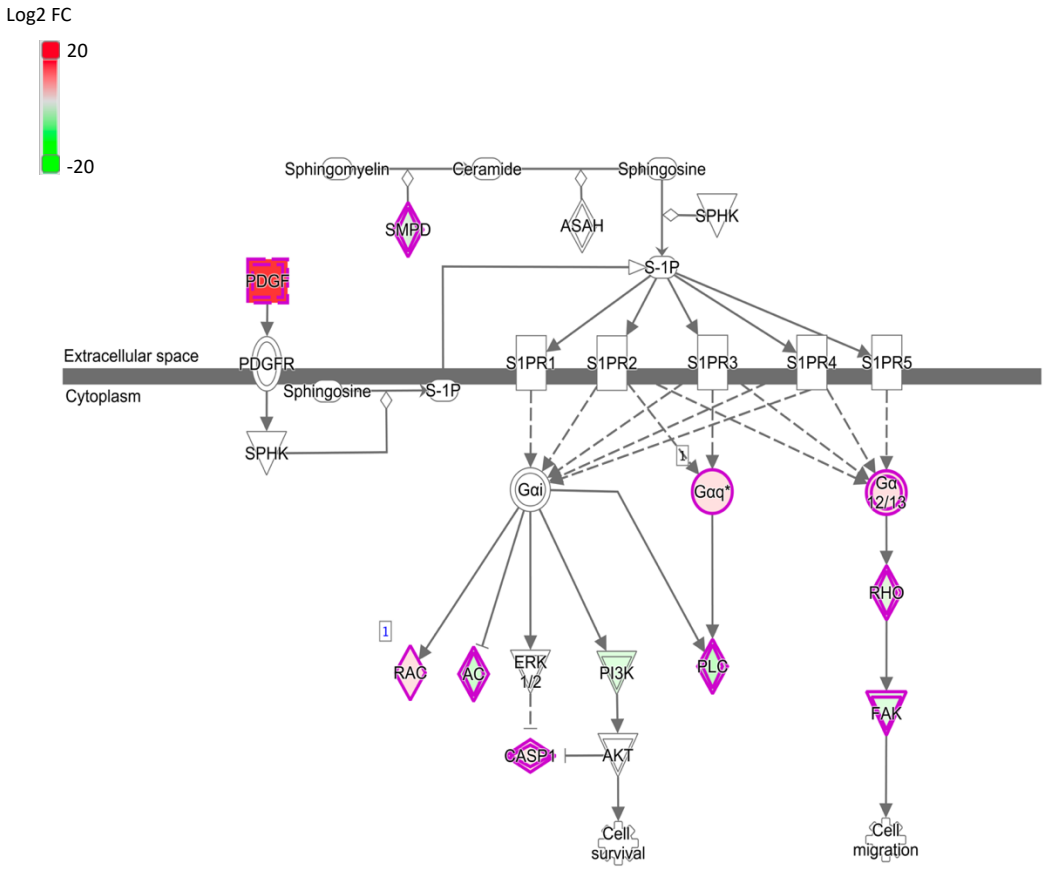


Figure 3.1 Canonical pathway 'Sphingosine-1-phosphate (S1P) signalling'.

Significantly enriched for genes whose expression correlates with differentially expressed between *Gata6*-WT and -KOmye mice, defined by Ingenuity Pathway Analysis software. Dysregulated genes outlined in pink; Up- and down-regulated genes shaded in red and green, respectively. The intensity of the node colour indicates the degree of regulation (-up or -down). The node shapes depict enzymes (◇), such as phosphatases (△), kinases (▽); transmembrane receptor (0); cytokines (☐); and other (○).

b) Lysosome fusion and function:

The canonical pathways involved in lysosome fusion and function were (Table 3.1.), grouped by biological processes, and listed below.

i) Endocytosis

Macropinocytosis Signalling (Figure 3.2)- refers to the actin-dependent formation of large endocytic vesicles of irregular size and shape called macropinosomes[163].

Clathrin-mediated Endocytosis Signalling (Figure 3.3)- is the major pathway for the internalization of nutrients, hormones and other signalling molecules from the plasma membrane into intracellular compartments[164].

Caveolar-mediated Endocytosis Signalling (Figure 3.4)- plays a role in various cellular processes such as endocytosis, cellular signalling and lipid recycling. Caveolae are cholesterol- and SL-rich, flask-shaped invaginations of the plasma membrane, which are involved in the internalization of membrane components, extracellular ligands, viruses and bacterial toxins[165].

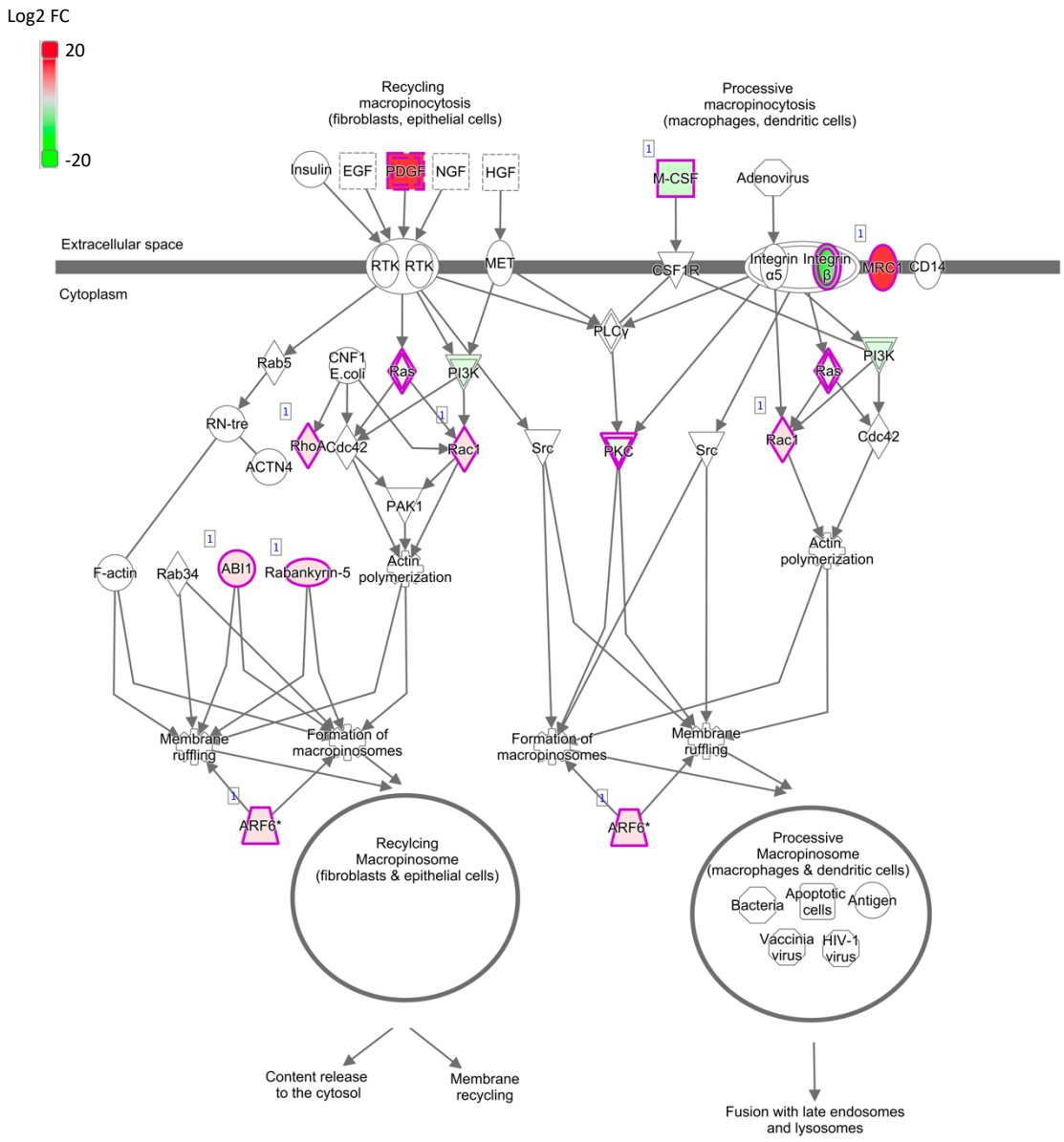


Figure 3.2 Canonical pathway 'Macropinocytosis Signalling'.

Significantly enriched for genes whose expression correlates with differentially expressed between *Gata6*-WT and -KO^{mye} mice, defined by Ingenuity Pathway Analysis software. Dysregulated genes outlined in pink; Up- and down-regulated genes shaded in red and green, respectively. The intensity of the node colour indicates the degree of regulation (-up or -down). The node shapes depict enzymes (◇), such as phosphatases (△), kinases (▽); transmembrane receptor (0); cytokines (□); and other (○).

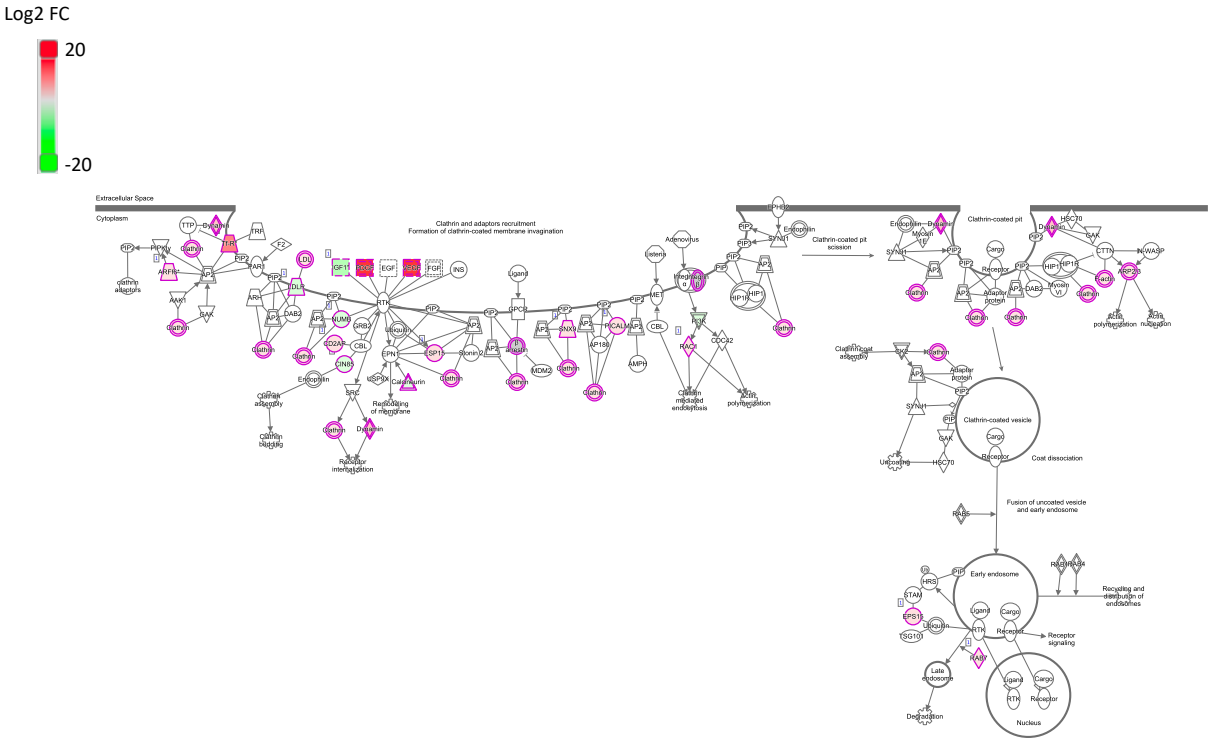


Figure 3.3 Canonical pathway 'Clathrin-mediated Endocytosis Signalling'.

Significantly enriched for genes whose expression correlates with differentially expressed between *Gata6*-WT and -KOmye mice, defined by Ingenuity Pathway Analysis software. Dysregulated genes outlined in pink; Up- and down-regulated genes shaded in red and green, respectively. The intensity of the node colour indicates the degree of regulation (-up or -down). The node shapes depict enzymes (◇), such as phosphatases (△), kinases (▽); transmembrane receptor (0); cytokines (□); and other (○).

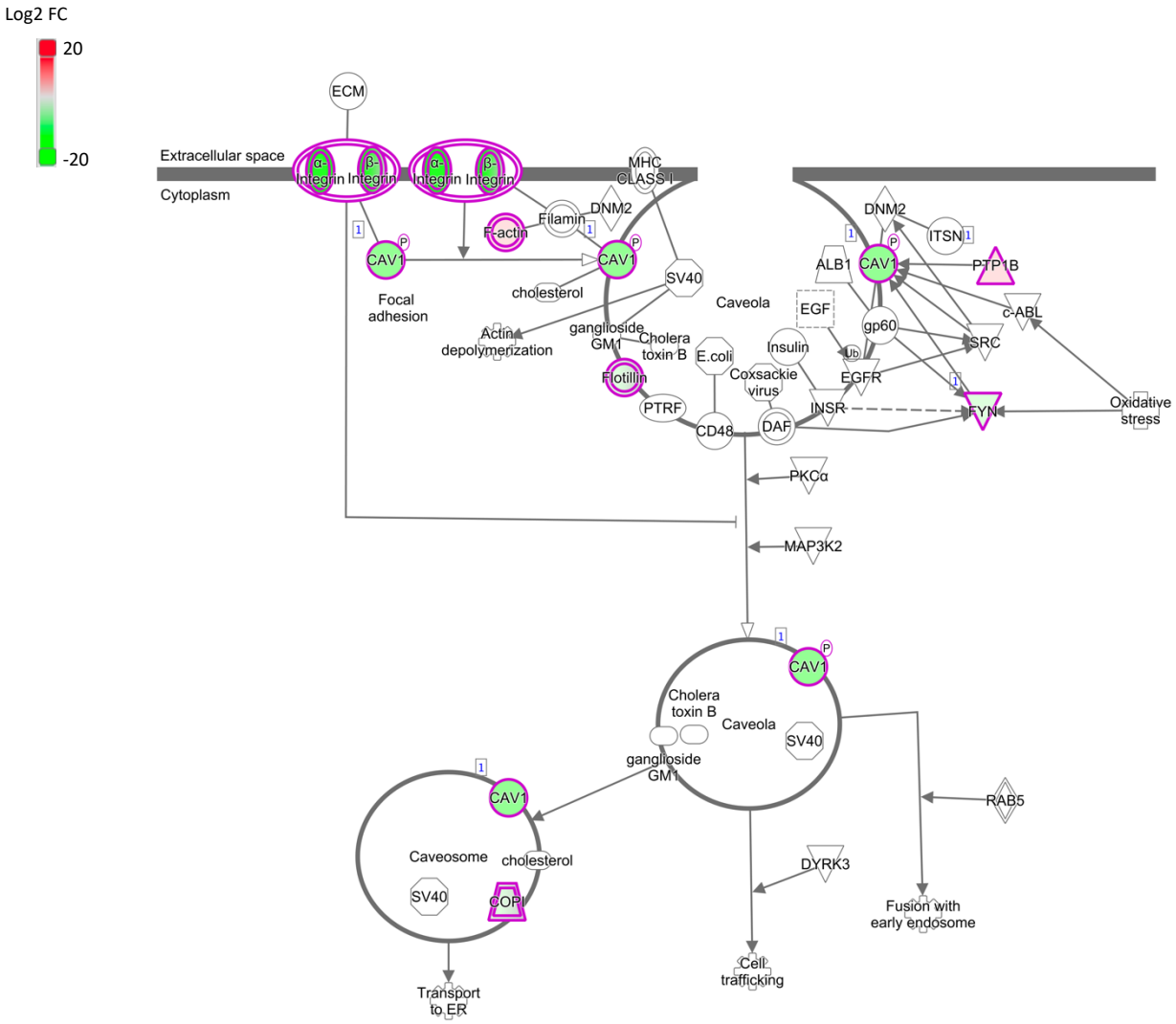


Figure 3.4 Canonical pathway 'Caveolar-mediated Endocytosis Signalling'.

Significantly enriched for genes whose expression correlates with differentially expressed between *Gata6*-WT and -KO^{mye} mice, defined by Ingenuity Pathway Analysis software. Dysregulated genes outlined in pink; Up- and down-regulated genes shaded in red and green, respectively. The intensity of the node colour indicates the degree of regulation (-up or -down). The node shapes depict enzymes (◇), such as phosphatases (△), kinases (▽); transmembrane receptor (0); cytokines (☐); and other (○).

ii) Phagocytosis

Phagosome formation (Figure 3.5)- A phagosome is formed by the fusion of the cell membrane. Phagosomes have membrane-bound proteins to recruit and fuse with lysosomes to form mature phagolysosomes[166].

Phagosome maturation (Figure 3.6)– As the phagosome matures, it becomes more acidic from pH 6.5 to pH 4 and gains characteristic protein markers and hydrolytic enzymes. The different enzymes function at different optimal pH, forming a range so they each work in narrow stages of the maturation process[166].

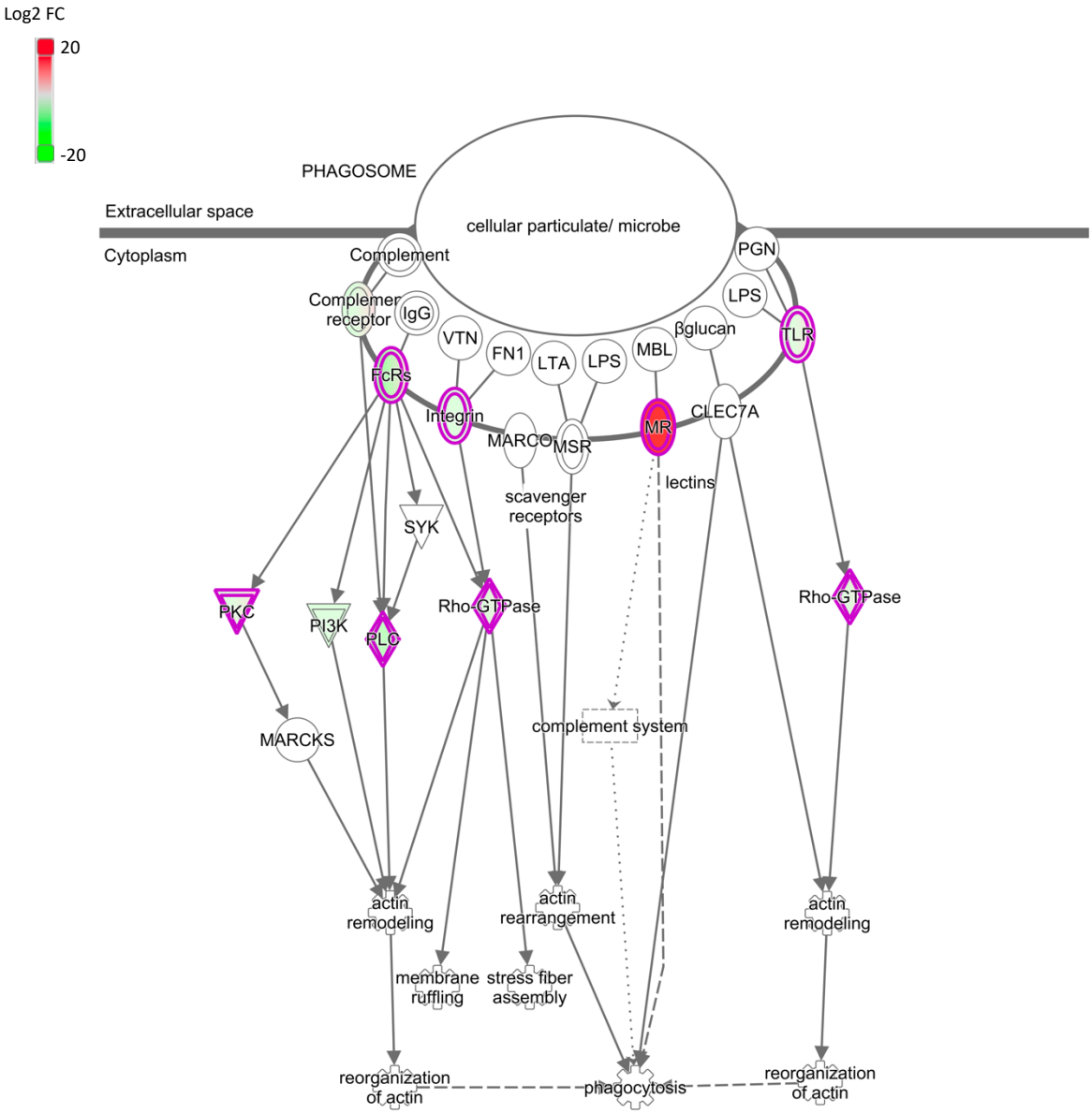


Figure 3.5 Canonical pathway 'Phagosome formation'.

Significantly enriched for genes whose expression correlates with differentially expressed between *Gata6*-WT and -KO^{mye} mice, defined by Ingenuity Pathway Analysis software. Dysregulated genes outlined in pink; Up- and down-regulated genes shaded in red and green, respectively. The intensity of the node colour indicates the degree of regulation (-up or -down). The node shapes depict enzymes (◇), such as phosphatases (△), kinases (▽); transmembrane receptor (0); cytokines (□); and other (○).

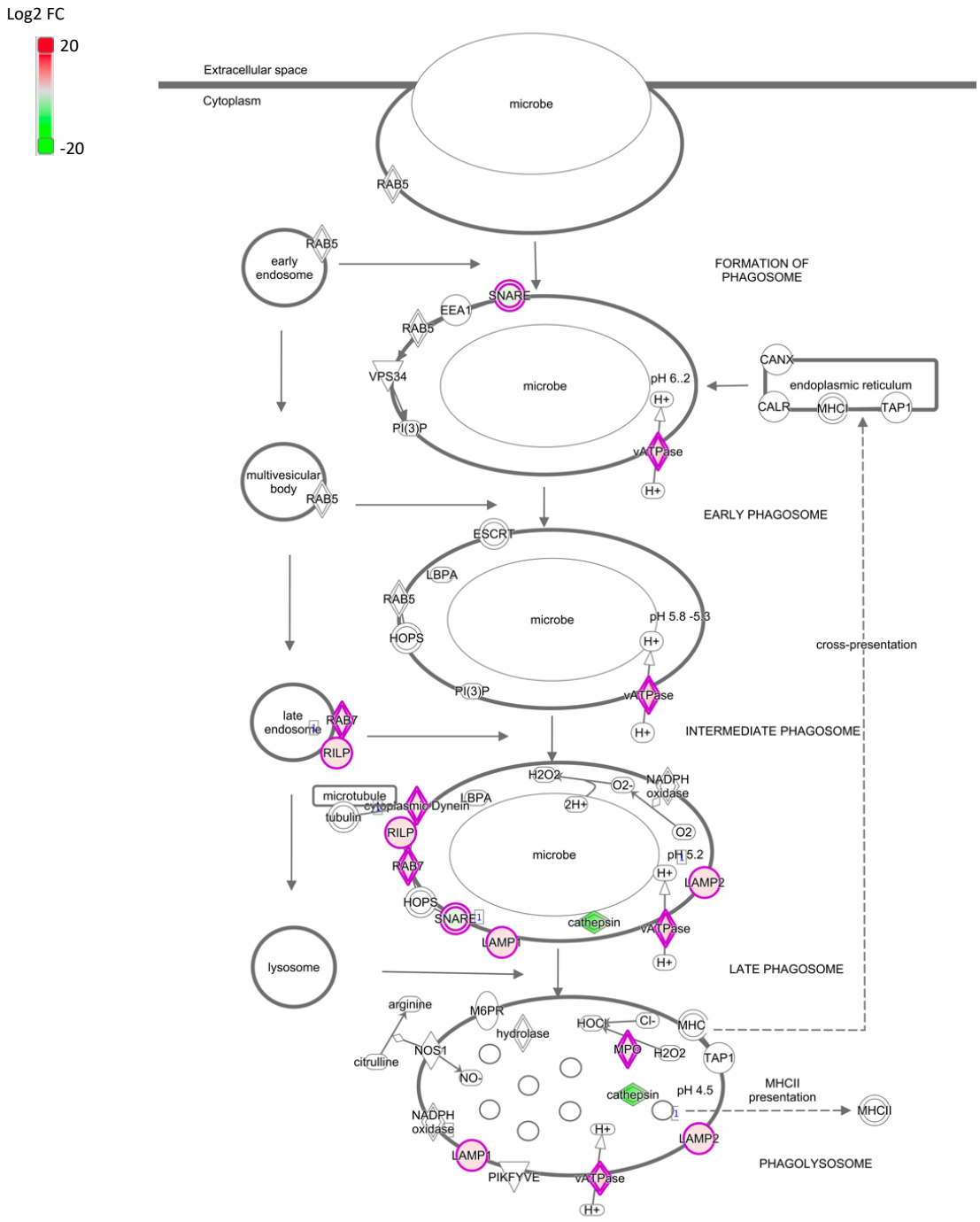


Figure 3.6 Canonical pathway 'Phagosome maturation'.

Significantly enriched for genes whose expression correlates with differentially expressed between *Gata6*-WT and -KO^{mye} mice, defined by Ingenuity Pathway Analysis software. Dysregulated genes outlined in pink; Up- and down-regulated genes shaded in red and green, respectively. The intensity of the node colour indicates the degree of regulation (-up or -down). The node shapes depict enzymes (◇), such as phosphatases (△), kinases (▽); transmembrane receptor (0); cytokines (□); and other (○).

3.2.2 Mechanistic Networks: inferring likely causal mechanisms

The mechanistic networks algorithm in IPA was used to determine which functions/mechanisms directly related to lipid metabolism had been altered as a consequence of loss of GATA6. The algorithm uses the network edges between pre-determined upstream regulators for which there is statistical evidence (P-value < 0.05) that the corresponding relationship is likely relevant for the causal mechanism behind the dataset. The most significant causal edges between regulators are then used to construct networks downstream of a 'master' regulator in order to indicate potential causal mechanisms [167].

The algorithm generated networks comprised of the genes present in the dataset and assigned a Z-score. The score is a numerical value that considers the number of focus genes and the size of the network. It is used to rank networks according to their relevance in regards to the genes in the input dataset, but it does not necessarily correlate with the biological significance of the network[160]. Using this process, I found 68 causal mechanistic categories with p-values<0.05; with the Lipid metabolism category (p-value = 2.15×10^{-08} - 2.42×10^{-05}), including 14 networks comprising of 202 genes (Table 3.2). This result pointed to the possibility that GATA6 deficiency leads to dramatic changes in lipid metabolism. Thus, the next step was to pinpoint which particular lipid families were the most affected by GATA6 deficiency by identifying causal networks that infer similar mechanisms/effects. In the next section I will select genes related to lysosomal metabolism and investigate whether the disruption of lysosomal function could contribute to lipidomic changes in GATA6-deficient cells.

Table 3.2 Lipid metabolism related networks significantly altered by *Gata6* expression.

Diseases or Functions Annotation	p-Value	Predicted Activation State	Activation z-score	Molecules	# Molecules
Fatty acid metabolism	2.17E-08		-0.579	ABCA1,ABCC1,ABCG1,ACACA,ACAT1,ACOT9,ACSS1,ALOX15,AQP9,ARSA,ATP8A1,B4GALT6,C5AR1,CASP8,CAV1,CAV2,CD82,CD9,CERK,COL4A3BP,COTL1,CPT1A,CS,CSF1,CXCL12,DBI,FABP4,FAM213B,FAR1,FASN,FCGR1A,FGR,FPR1,FYN,GBA,GDF15,Gk,GLTP,HACD3,HMOX1,HSD17B4,IGF1,ITGB3,KCNN4,KITLG,LAMTOR1,LDLR,LIPA,LPL,LTA4H,LTC4S,MAP3K8,MDH1,ME2,MECR,MIF,MITF,MYO5A,NCF1,NPC1,NPC2,OLR1,OSBPL2,P2RX1,PARK7,PDHA1,PLTP,PRELID1,PRKCD,PTGIS,PTGR1,PTGS1,PTK2B,RAC1,RARB,RELA,RHOA,RTN4,SCARB1,SDHA,SEC13,SGPP1,SLC36A1,SLC9A3R1,SLC9A3R2,SMPD1,SMPD2,SOD1,SPTLC2,ST3GAL5,TLR4,TRIAP1,TSPO	93
Concentration of phospholipid	5.85E-08		0.923	ABCA1,ABCG1,ABHD12,ACACA,ANXA5,ARF6,BAX,BSCL2,CAV1,CD38,COL4A3BP,CSF1,DAB2IP,FASN,GNAQ,GOLPH3,HINT1,INPPL1,KITLG,LDLR,LPL,LSR,MBOAT7,NPC1,PLTP,PRKCD,PSEN2,PTEN,RAB27A,RHOA,RILP,ROCK2,SCARB1,SIK3,SMPD1,SMPD2,SPTLC2,STARD4	38
Synthesis of lipid	3.84E-07	Decreased	-2.297	ABCA1,ABCC1,ABCG1,ACACA,ACAT1,ACAT2,ACSS1,ADORA2B,ALOX15,ARF6,ARSA,ASPA,B4GALT6,BSCL2,C5AR1,CASP8,CAV1,CAV2,CD82,CD9,CERK,CNBP,COTL1,CSF1,CXCL12,CYP26A1,CYP7B1,DBI,DENND1A,DPM2,FAM213B,FAR1,FASN,FCGR1A,FGR,FPR1,FYN,GBA,GDF15,HACD3,HDC,HMOX1,IGF1,IGFBP4,INPPL1,ITGA6,ITGB1,ITGB2,KITLG,LDLR,LPCAT1,LPCAT2,LPL,LTA4H,LTC4S,MAN2B1,MAP3K8,ME2,MGAT2,MIF,MITF,MYO5A,NCF1,NPC1,NPC2,NR3C1,OLR1,P2RX1,PARK7,PDHA1,PLA2G16,PLTP,PRKCD,PTGIS,PTGS1,PTK2B,RAB27A,RAC1,RARB,RELA,RGS2,RGS3,RHOA,RHOQ,SCARB1,Sf1,SGPP1,SH3KBP1,SLC6A6,SLC9A3R1,SLC9A3R2,SMPD1,SMPD2,SOD1,SORBS1,SPTLC2,ST3GAL5,STARD4,TAZ,TBXAS1,TCF7L2,TLR4,TSPO	103
Metabolism of membrane lipid derivative	1.03E-06		-1.585	ABCA1,ABCC1,ABCG1,ALOX15,APLP2,ARF6,ARSA,ASPA,B4GALT6,BAX,CASP8,CAV1,CD82,CD9,CERK,CLN8,CNBP,COL4A3BP,CREM,CYP7B1,DBI,DPM2,FAR1,FASN,GATA6,GBA,HMOX1,IGF1,INPPL1,ITGB1,ITGB2,KITLG,LDLR,LIPA,LPCAT1,LPCAT2,LPL,MBOAT7,NPC1,ORMDL2,P2RX1,PLA2G16,PLA2G7,PLEK,PLTP,PRKCD,PTEN,RAB27A,RAC1,RHOA,RHOQ,SCARB1,SERINC3,SGPP1,SH3KBP1,SLC6A6,SMPD1,SMPD2,SPTLC2,ST3GAL5,TAZ	61
Accumulation of cholesterol	1.07E-06	Decreased	-3.241	ABCA1,ACAT1,ACAT2,ACP2,CAV1,CREM,FABP4,GSK3A,HMOX1,LAMP1,LAMP2,LDLR,NPC1,NPC2,PLTP,SMPD1	16
Accumulation of sterol	1.33E-06	Decreased	-3.313	ABCA1,ACAT1,ACAT2,ACP2,CAV1,CREM,FABP4,GSK3A,HMOX1,LAMP1,LAMP2,LDLR,NPC1,NPC2,PLTP,SCARB1,SMPD1	17
Concentration of lipid	1.97E-06		0.886	ABCA1,ABCG1,ABHD12,ACACA,ACAT1,ACAT2,ACSS1,ADIPOR1,ADIPOR2,ADORA2B,ALOX15,ANXA5,AQP9,ARF6,ARNTL,BAX,BSCL2,CASP8,CAV1,CCDC80,CD38,CDKN2C,CEACAM1,CERK,COL4A3BP,COMT,CPT1A,CSF1,CXCL12,CYP26A1,DAB2IP,DBI,FABP4,FASN,GATA6,GBA,GFRA2,Gk,GNAQ,GOLPH3,GSTK1,H2AFY,HINT1,Hmgn3,HMOX1,HP,HSD17B4,IDH1,IGF1,IL18BP,INPPL1,ITGB3,KITLG,KPNA1,LDLR,LIPA,LPL,LRPAP1,LSR,LTA4H,LTC4S,MAP3K8,MBOAT7,MGAT4A,MIF,MRC1,NCOA1,NEU1,NPC1,NPC2,NR3C1,OLR1,PDE10A,PDE3B,PEX11A,PHYH,PLTP,PMP22,PPP3CA,PRKCD,PSEN2,PTEN,PTGIS,PTGS1,PTPN1,RAB27A,RAB7A,RAC1,RGCC,RHOA,RILP,ROCK2,SCARB1,SGPP1,SIK3,SLC4A4,SLC9A3R1,SMPD1,SMPD2,SOD1,SOD2,SORBS1,SPTLC2,SRGAP3,ST3GAL5,STARD4,TAZ,TCF7L2,TGFBR2,TLR4,VAMP3,ZBTB16	112
Accumulation of lipid	4.48E-06	Decreased	-2.884	ABCA1,ABCC1,ACACA,ACAT1,ACAT2,ACP2,ADIPOR1,ADIPOR2,APOBR,ARSA,BAX,BRD8,CAV1,COL4A3BP,CPT1A,CREM,CSF2RB,CXCL12,DBI,EPAS1,FABP4,FASN,FGR,GBA,GSK3A,H2AFZ,HCK,HMOX1,IDH1,IGF1,LAMP1,LAMP2,LDLR,LPL,MIF,NPC1,NPC2,PLTP,PTEN,PTPN1,RENBP,SCARB1,SCARB2,SHTN1,SLC36A1,SMPD1,SMPD2,TLR4	48
Accumulation of steroid	5.90E-06	Decreased	-3.447	ABCA1,ACAT1,ACAT2,ACP2,CAV1,CREM,FABP4,GSK3A,HMOX1,IGF1,LAMP1,LAMP2,LDLR,NPC1,NPC2,PLTP,SCARB1,SMPD1	18
Quantity of sphingolipid	1.16E-05		0.981	ARF6,CASP8,CAV1,CERK,COL4A3BP,CSF1,GBA,HINT1,IGF1,NPC1,PLTP,PMP22,PRKCD,PSEN2,RAC1,RHOA,SCARB1,SGPP1,SMPD1,SMPD2,SPTLC2	21
Quantity of glycosphingolipid	1.46E-05		0.855	ARF6,CASP8,CAV1,COL4A3BP,CSF1,GBA,HINT1,IGF1,NPC1,PLTP,PMP22,PRKCD,PSEN2,RAC1,RHOA,SCARB1,SGPP1,SMPD1,SMPD2,SPTLC2	20
Storage of lipid	1.99E-05		-0.286	ABCA1,ABCG1,ARSA,BSCL2,CAV1,CISD1,EHD1,FABP4,ITGB3,LIPA,LPL,MGAT2,NPC1,NRIP1,OSBPL11,SCARB1,SMPD1	17
Esterification of cholesterol	2.21E-05		-0.080	ABCA1,ABCG1,ACAT1,ACAT2,CAV1,LAMTOR1,LDLR,LPL,OSBPL2,SCARB1,STARD4	11
Concentration of sphingomyelin	2.49E-05		1.998	ARF6,CAV1,COL4A3BP,HINT1,PLTP,PRKCD,PSEN2,SMPD1,SMPD2,SPTLC2	10

3.2.3 Lysosomal Pathways and genes

Regulation of lipid metabolism is key to the homeostasis of cells and organisms, and the lysosome is an important element of this network, due to its capacity to process and sort exogenous and endogenous lipids[152]. The results obtained in section 3.2.1 Identified four canonical pathways involved in lysosome fusion and function, therefore supporting the hypothesis that the lysosomal function is regulated by the GATA6 transcription factor in pResMϕ.

In order to explore the possibility that GATA6 regulates lysosomal function, I first selected a group of 566 genes known to be involved in lysosome biology (which I names LBG dataset). Next, I used the expression values for these genes to perform a guided analysis of the changes occurring in *Gata6*-KO^{mye} pResMϕ. Given the impact of sample size in p-value adjustment, it was necessary to reassess the significance of this data subset, by correcting the p-values using the Benjamini-Hochberg method. Thus, I found that of the LBG dataset, 101 genes are statistically different (p value<0.05). Compared to the values used in section 3.2.1 there were now 31 new genes considered relevant to use in the analysis of lysosomal mechanisms and pathways. The volcano plot depicted in Figure 3.7 shows how the genes from the LBG dataset are regulated in *Gata6*-KO^{mye} pResMϕ. A total of 33 genes with significantly altered expression levels were found (at least 2-fold difference and with p-value < 0.05), with 12 and 21 genes either down or up regulated, respectively.

The LBG dataset was analysed using statistical gene-set-enrichment methods where differentially-expressed genes are interrogated using lists of genes associated with a particular biological function or pathway. Significantly differentially expressed genes (corrected p-value<0.05 and fold change≥ 1.5) were uploaded and analysed in IPA. I first determined which canonical pathways were relevantly changed, and I found that the results obtained via this approach supported those in Section 3.2.1, where the combination of the most enriched canonical pathways, related to lipid metabolism, suggested a dysregulation of lysosome function, but also provided further insight regarding which lipid pathways/classes could be the most sensitive to the deletion of *Gata6*. Table 3.3

lists the canonical pathways related to lipid metabolism of the LBG dataset that were significantly enriched in *Gata6*-KO^{mye} pResMϕ. Interestingly, five of the eight pathways were directly linked to SL metabolism-related genes. Thus, my results suggested that GATA6 deficiency may lead to dramatic dysregulation of SL metabolism.

3.2.4 SL metabolism

A closer inspection into which functions/mechanisms, directly related to lipid metabolism, were altered as a consequence of GATA6 deficiency, revealed that the networks 'quantity of sphingolipid', 'concentration of sphingomyelin' and 'quantity of glycosphingolipid', (Figure 3.8) were all activated, suggesting a pattern of network alterations that led to accumulation of SL. The generated networks are presented as a graph indicating the molecular relationships between genes and gene products. Genes are depicted as nodes, and the biological relationship between two nodes is represented as an edge (line). The intensity of the node colour indicates the degree of regulation (-up or -down). The node shapes depict enzymes (◇), such as phosphatases (△), kinases (▽); transmembrane receptor (0); cytokines (□); and other (○).

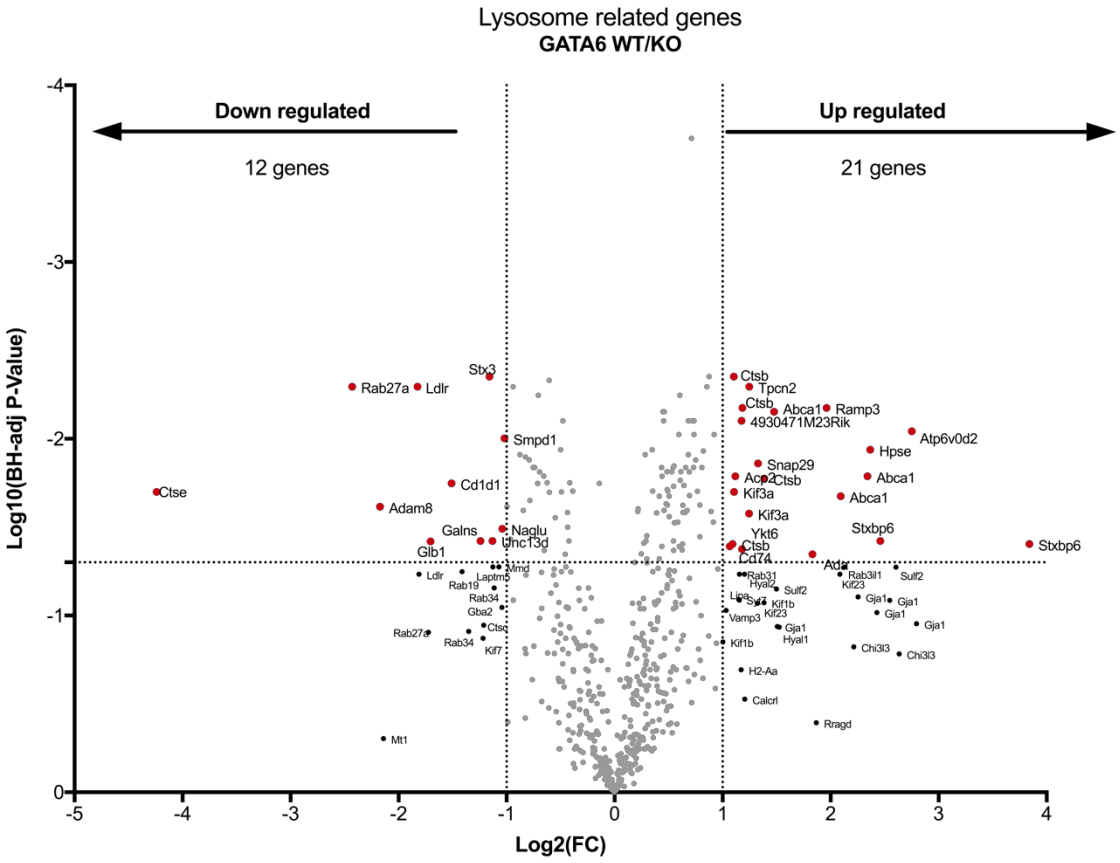


Figure 3.7 GATA6 deficiency leads to dysregulation of lysosomal metabolism related genes.

'Volcano plot' of statistical significance against fold-change between *Gata6*-WT and *Gata6*-KO^{mye} pResMφ, demonstrating the most significantly differentially expressed genes. Notably 2-fold down-regulated and up- regulated genes with p-value<0.05 are indicated. A total of 36 genes with significantly altered (red dots) expression levels were found (at least 2-fold altered and with p-value < 0.05), with 12 and 24 genes either down or up regulated, respectively. Statistical significance was assessed using Benjamini-Hochberg method.

Table 3.3 - The most significantly enriched lipid metabolism related pathways in the LBG dataset for *Gata6*-KO pResMϕ according to IPA[2].

Ingenuity Canonical Pathways	-log(p-value)	Ratio	z-score
Phagosome maturation	33.6	0.25	NaN
Autophagy	23.2	0.35	NaN
Ceramide Degradation	4.23	0.5	NaN
Clathrin-mediated Endocytosis Signaling	3.8	0.0558	NaN
Sphingosine and Sphingosine-1-phosphate Metabolism	3.79	0.375	NaN
Sphingomyelin Metabolism	3.79	0.375	NaN
Sphingosine-1-phosphate Signaling	3.37	0.0656	-1.342
Ceramide Signaling	1.93	0.0538	NaN

[2] A positive z-score indicates a predicted activation, and a negative z-score indicates a predicted inactivation of the enriched pathway. NaN, a z-score cannot be calculated for all Ingenuity canonical pathways.

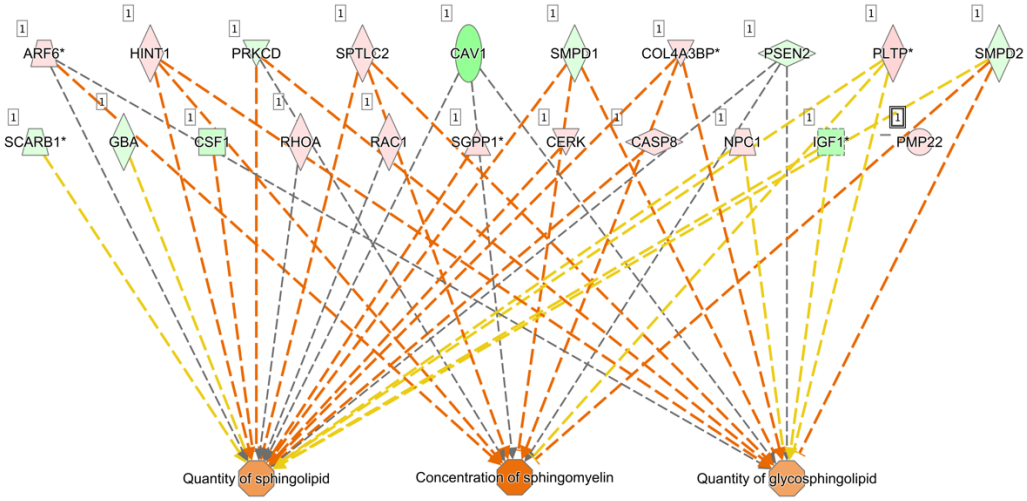


Figure 3.8 Causal networks infer SL accumulation in GATA deficient ResMϕ.

Predicted regulatory networks from Causal Network Analysis and the impact on biological function. The network is graphically represented as nodes (genes) and edges (the biological relationship between genes). Genes are depicted as nodes, and the biological relationship between two nodes is represented as an edge (line). The intensity of the node colour indicates the degree of regulation (-up or -down). The node shapes depict enzymes (◇), phosphatases (△), kinases (▽), transmembrane receptor (0), cytokines (□), and other (○). Red and green shaded nodes represent up- and downregulated genes, respectively. Orange lines indicate that the regulation of that gene leads to activation, Yellow lines indicate that the regulation observed is inconsistent with expectations, while grey arrows indicate lack of pre-existing data to formulate expectations.

3.3 Discussion

The purpose of this chapter was to investigate whether the expression of genes related to lipid metabolism was regulated by the transcription factor GATA6 in pResMϕ. First, the microarray data from WT and *Gata6*-KO^{mye} mice, obtained by Prof. Phil Taylor's group was analysed by employing the canonical pathways feature in IPA in order to ascertain the probable downstream effects of *Gata6* KO on cellular biology. This analysis was carried out with particular emphasis on lipid related pathways by focusing not only on pathways and causal networks that could be involved in lipid metabolism, but also on those that could affect lipid levels by altering lipid catabolism, transport, and storage.

Using this method, I identified eight canonical pathways that suggested how lipid metabolism was regulated by GATA6 in pResMϕ. These can be divided into two main sets, (i) regulation of specific lipid categories, and (ii) lysosome fusion and function. The first comprised three lipid categories, phospholipids, fatty acids and SL, whilst the latter covered the different routes of endocytosis as well as crucial steps in phagosome formation. This indicates that *Gata6* expression is critical for lipid biosynthesis and accumulation, in particular, by regulating SL degradation by lysosomes. That in turn could lead to lysosomal storage disorders. Thus, my results provide possible link between *Gata6* expression and downstream signalling related to lysosomal dysfunction.

There are over 40 lysosomal storage disorders known, including several sphindolipidoses where an intra-lysosomal accumulation of SL occurs, with the most well studied of these being Niemann-Pick disease[125]. Additionally, it was recently shown that a marked decrease in cathepsin E (CTSE), a proteinase mainly localized on the endosomal compartment in Mϕ[168], and highly down-regulated in GATA6-deficient cells, can lead to a novel form of lysosome storage disorder[169] and lead to abnormal lipid metabolism[170]. Interestingly, the microarray data showed a downregulation of the *Ctse* gene, indicating that CTSE could be involved in maintaining lipid homeostasis in Mϕ.

The mechanistic network algorithm from IPA was then used to infer which functions/mechanisms were likely to lead to changes in lipid metabolism as a consequence of GATA6 deficiency. Thus, only the hypotheses that added relevant information were considered (*i.e.* lipid metabolism related causal networks), revealing 14 mechanistic networks of interest (Table 3.2). However, some of these data were puzzling, for instance, 'quantity of glycosphingolipid', 'quantity of sphingolipid' and 'quantity of sphingomyelin' had positive z-scores, indicating network activation; whilst at the same time, 'accumulation of lipid', 'lipid storage' and 'synthesis of lipid' had negative z-scores, suggesting a decreased activation state. It is possible though that this could indicate an increase in SL, combined with a decrease of overall lipid levels. Thus, taken together these analyses indicated that GATA6 signalling was involved in lipid metabolism in pResMϕ. These observations led to a further in-depth analysis of a data subset comprising of genes involved in lysosomal metabolism. Through this approach I found five canonical pathways involved in SL metabolism that were significantly enriched in GATA6-deficient pResMϕ (Table 3.3). The analysis of this data subset using the causal analysis algorithm revealed a particular pattern in gene regulation/network activation that ultimately could lead to an increase in SL levels (*e.g*. down-regulation of genes that lead to lipid consumption, up-regulation of genes that lead to lipid synthesis). SL are located in the outer leaflet of the plasma membrane and interact with cholesterol-containing specific membrane domains known as 'lipid rafts'. These are small regions (10-200 nm) that play a key role in the compartmentalisation of cellular processes by stabilising protein-protein or protein-lipid interactions. Further analysis of IPA data showed a dysregulation of other genes involved in Caveolin-mediated endocytosis. The presence of Caveolin-1 in a lipid raft creates an invagination of the membrane recognised as 'caveolae'[171]. This is of particular relevance as caveolae-mediated endocytosis functions as a 'true uptake mechanism' parallel to the clathrin mediated pathway. However, caveolar endocytosis provides a more selective and highly regulated way for uptake of specified substances[172]. Caveolae are well-studied in certain cell types (*e.g.*, endothelial cells, fibroblasts), it is thought to be absent in others (*e.g.*, lymphocytes). The presence of caveolae in Mϕ is less well-defined and indeed controversial, varying by Mϕ type[173]. Even

though, the precise function of caveolins in Mϕ lipid metabolism is still unclear, studies suggest their involvement in cholesterol transport to the membrane[174]. Additionally, caveolins may also be involved in the modulation of Mϕ immune responses, as well as, in the clearance of apoptotic cells in lesions[174].

In this Chapter, I performed analysis of gene expression in *Gata6*-KO^{mye} pResMϕ. I found dysregulated gene networks directly linked to lipid metabolism that can potentially impact the regulation of pResMϕ phenotype. In particular, the IPA analyses indicated enrichment in gene networks associated with an increase in the levels of SL. This class of lipids has key roles in the regulation of both trafficking and functions of immune cells[175], and their accumulation leads to sphingolipidoses, a group of inherited human diseases in which SL accumulate in one or more organs due to a degradation disorder[176]. The activation of gene networks indicating SL accumulation can potentially lead to changes in lipid levels not only in SL but in the other classes as well. It has previously been shown that perturbations in genes involved in SL metabolism, can lead to a remarkable heterogeneity in lipid states, affecting both SL and GPL[98]. Thus, my analysis identified one specific class of lipids to be directly linked to *Gata6* expression in pResMϕ (SL).

In conclusion, the findings in this Chapter showed that: 1) GATA6 transcription factor regulates the expression of lipid metabolism related genes in pResMϕ; 2) in particular, SL metabolism-related genes are affected by GATA6 deficiency, as evidenced by the suggested dysregulation of SL metabolism-related pathways and causal mechanisms; 3) a wide range of genes with roles in lysosomal function have been found to be dysregulated as a consequence of GATA6 loss. Hence, the first step in answering the question of how GATA6 affects lipid levels in pResMϕ, will be to establish an untargeted LC-MS lipidomics workflow. This will enable me to perform a global lipidomics analysis in *Gata6*-KO^{mye} and WT pResMϕ, and thoroughly characterise the impact of gene dysregulation on global lipid levels.

Chapter 4 Optimisation of LC-MS conditions for untargeted lipidomics of Mϕ-like cells

4.1 Introduction

4.1.1 Untargeted lipidomics

Studies conducted in Chapter 3 demonstrated that the GATA6 transcription factor regulates lipid metabolism related genes in pResMϕ. This work now needs to be extended through characterising how GATA6 regulates the levels of lipid metabolites themselves. Herein I will optimise and validate an untargeted LC-MS lipidomics workflow, that will later be used to characterise the lipidomic differences between *Gata6*-WT and -KO^{mye} pResMϕ.

Lipids are important structural components of cells, that also act as sources of energy and signalling mediators. Alterations in lipid metabolism are involved in the onset and progression of a variety of diseases such as cancer, cardiovascular disorders, neurodegenerative diseases, obesity and diabetes[177–181]. Hence the increasing interest in the comprehensive characterisation of the lipidome (lipidomics), using high-resolution mass spectrometry. There are numerous lipidomics methods available. However, the structural diversity of lipids makes getting a complete overview of all lipids in a system extremely challenging[182]. High-resolution lipidomics approaches able to profile lipids in a complex matrix, and remove artefacts (*i.e*. adducts and isotopes), are not widely available, thus estimating the total lipidome is extremely difficult[93].

The most common approach to this problem until recently was a comprehensive lipidomic analysis by 'separation simplification'[153]. This approach uses several different methods that each address the structural and biophysical diversity of lipid categories. In that way, the extraction, LC separations and MS protocols are separately optimised to obtain quantitative information about the compounds in each lipid category separately, in a targeted fashion[182]. The drawback of targeted lipidomics is that only known species of lipids are measured[93].

Conversely, untargeted approaches are used for global profiling of lipids in biological systems without bias. While several analytical approaches can be used to perform untargeted lipidomics, liquid chromatography/mass spectrometry

(LC/MS) is often the technique of choice, due to the large number of metabolites that can be measured in a single analysis[183]. Using LC/MS, tens of thousands of "features" can be analysed in a single lipid extract (where a feature is defined as an ion with a unique *m/z* and retention time). At the cost of increased run time, LC separation reduces the complexity of the matrix introduced into the mass spectrometer, thus decreasing ion-suppression during ionization. As a result, low abundance species are more readily detected[184]. With untargeted lipidomics the aim is often to determine which of these features are dysregulated between two or more sample sets.

Recently, an untargeted method for mapping the platelet lipidome was published using a lipidomic workflow that maximizes coverage, ensuring that low-abundance lipids were detected and verified, essential step for the discovery of signalling lipids[93]. For these reasons, this method was chosen as a starting point for the lipidomic analysis of Mϕ populations of interest in my project. However, this workflow had been optimised for the analysis of platelet lipids where a higher number of cells was available, than I could easily obtain using *Gata6*-WT and -KO^{mye} pResMϕ. Thus, herein I will use Mϕ-like cells to optimise and validate an untargeted lipidomics approach that can later be used to analyse pResMϕ.

Studies in the past 2 decades have provided remarkable insight into the functional and phenotypic diversity of Mϕ, which reflects their vital role in innate immunity, homeostasis and pathogenesis[185]. The majority of Mϕ *ex vivo* studies utilize only a few Mϕ phenotypes for practical reasons[111] (as discussed below).

The peritoneal cavity provides an easy-access source of pResMϕ, which can be isolated through lavage with PBS, yielding *c.a.* 1 million cells per mouse[186]. This is a relatively low yield that makes large scale experiments quite difficult. Alternatively, thioglycollate-elicited Mϕ (TGEM) can be isolated with much higher yields (30-40 million cells/mouse), however their phenotype is significantly different from that of pResMϕ as they are inflammatory Mϕ derived from monocytes[187]. BMDM are another source of primary Mϕ. Nevertheless, BMDMs require isolation of stem cells that are further cultured for 1 week in the

presence of M-CSF for differentiation, which represents a fairly artificial environment[111]. Lastly, RAW 264.7 cells are one of the most widely used Mϕ phenotypes used in cell culture due to their ability to replicate rapidly. RAW cells represent an immortalized phenotype derived from peritoneal tumours induced in BALB/c male mice by the Abelson murine leukaemia virus[111]. These cells have been passaged for over 20 years *ex vivo*, thus relating information from RAW cells to a biological context requires some direct comparison with primary Mϕ[111].

A previous attempt at characterizing the global murine Mϕ lipidome relied solely on the use of RAW cells[95] and did not address the issue of the different Mϕ phenotypes and how their lipid metabolism might be differentially regulated. Additionally, studies where the different Mϕ phenotypes were considered focused solely on FA metabolism[111,188]. In this Chapter I will optimise and validate an untargeted lipidomics approach, for global lipid profiling of murine pResMϕ and RAW cells, ensuring that the occurrence of false positive features is minimised.

This Chapter describes the modifications implemented to the workflow in order to adapt it to the samples/questions of interest for the lipidomics analysis of *Gata6*- WT and -KO^{mye} pResMϕ. I will use this workflow to profile the lipidome of murine pResMϕ and RAW cells to uncover cell-type specific differences in order to ensure that the methods employed are optimised for the untargeted lipidomic analysis of Mϕ.

4.1.2 Aims

- Studies described in this Chapter will establish an untargeted LC-MS method for the lipidomic analysis of Mϕ. Here, I will:
    - Evaluate the sensitivity of the untargeted lipidomics approach.
    - Check possible artefacts in data analysis.
    - Optimize the LTQ Orbitrap settings for optimal performance.
    - Analyse differences in lipid molecular species between primary and immortalised Mϕ cells used for ex vivo studies.
    - Establish a lipidomic analysis workflow that minimizes the occurrence of false positive and false negative features, providing a more manageable and comprehensive dataset.
    - Characterise the lipidomic differences between Mϕ-like cells.

4.2 Results

An untargeted lipidomics approach based on LC-MS will be used to profile lipid differences in murine Mϕ. However, there is no standardised workflow and researchers use a wide variety of high-resolution mass spectrometers, operating under different operating conditions. To this end, I first investigated the quality of lipidomics data obtained from control samples (peritoneal lavages) analysed under several conditions. LC-MS conditions and data processing methods were optimised in a systematic manner, in order to achieve a workflow capable of profiling lipidomic changes in murine Mϕ, and will be described in detail in the following sections.

4.2.1 Optimisation of LC-MS conditions

a) Mass range

In previous studies using similar conditions, two LC-MS runs were combined for untargeted lipidomics, providing coverage of both high and low *m/z* values (*m/z* 100 - 900 and *m/z* 800 - 1600)[93]. This ensured that the sensitivity of the method was not compromised. However, being able to combine these two ranges into one would mean reducing by half the number of runs needed and consequently the amount of sample required, which is paramount when dealing with animal samples, in particular with samples from *Gata6*-WT-KO^{mye} pResMϕ, due to their low cell yield.

To determine whether combination of both high and low mass ranges would result in loss of features/sensitivity, I analysed a lipid extract from Raw 264.7 cells in negative mode using *m/z* range from 100 – 1600 amu ("Full" range) and a combined range from 100 - 900 and 800 - 1600 *m/z* ("combined" range), where the same sample was run twice, scanning from 100 – 900 *m/z* and a second time with a 800 – 1600 *m/z* range. Following the LC-MS analysis the data from the 100 - 900 and 800 – 1600 *m/z* runs were combined. It is important to note that the data comparison was done using pre-processed data from an initial step in the data analysis and therefore, both the number of frames and peaks obtained was compared.

Results showed that the number of features was similar to both mass ranges ("combined" and "full"), 2281 and 2323 features for the "combined" and "full" ranges, respectively, revealing a 2 % increase in the number of features found in the full range, of which 73 % were common to both approaches (Figure 4.1). Thus, the data indicates that there is no significant loss of information, when comparing the two approaches.

a) Improvement of mass accuracy of high-resolution LC-MS data using background ions

Mass accuracy and mass reproducibility are key parameters of mass spectrometric performance and lipid species identification. Mass accuracy is commonly specified as the ratio of the m/z measurement error ($\Delta m/z$) to the true m/z. The mass accuracy is typically reported in parts per million (ppm, 10^{-6}) (Equation 4.1). Mass reproducibility is the standard error of the m/z measurement[189].

Equation 4.1 Mass accuracy

$$Mass\ accuracy\ (ppm) = \frac{observed\ m/z - theoretical\ m/z}{theoretical\ m/z} \times 10^6$$

Mass accuracy and mass reproducibility are linked to both the resolving power and the mass spectrometer hardware. Essentially, a relative mass reproducibility of ± 1 ppm means that the measurement error of a lipid ion with m/z 304.4669 will be within ± 0.0003 amu. Hence, the measured m/z value of the lipid ion will be in the range of m/z 304.4970 ± 0.0003 amu. Thus, the identification of lipid ions is usually performed using a "m/z tolerance window" (5 ppm max), set according to the mass resolution, mass accuracy and mass reproducibility of the acquired data. In order to obtain a high mass accuracy – and specific lipid identification – it is necessary to accurately calibrate the mass spectrometer. However, large lipidomic experiments with data acquisition over several days will be influenced by drifts of mass analyser accuracy, leading to a calibration m/z offset that will affect the mass calibration but not the mass reproducibility. One way of minimizing the adverse effects of m/z calibration drift is by internal calibration, using internal lipid standards or sample

contaminants with known *m/z* value that are present in all samples. This latter approach is called "lock mass" [190,191].

While the concept of a lock mass is well known it is not frequently employed. This is due, in part, to the effort of introducing a lock mass into the samples being analysed. However, herein the lock-masses used for positive and negative mode were from stable contaminants found in all samples. Before each analysis, the instrument was externally calibrated and at least three methanol blanks were analysed by LC-MS. Internal mass calibrants (lock-masses) were used in positive and negative ion mode and consisted of diisooctyl phthalate (*m/z* 391.2842 $[M-H]^-$) and sodium dodecyl sulfate (*m/z* 265.1479) respectively. Implementing this strategy resulted in a significant improvement in mass accuracy and reproducibility (Figure 4.2). Several lipid standards were analysed in order to assess the mass accuracy values obtained with and without lock-mass (Figure 4.3).

Having successfully optimised the LC-MS conditions, in the next section I will establish methods for data analysis and validate the untargeted lipidomics workflow using lipid extracts from Mϕ-like cells.

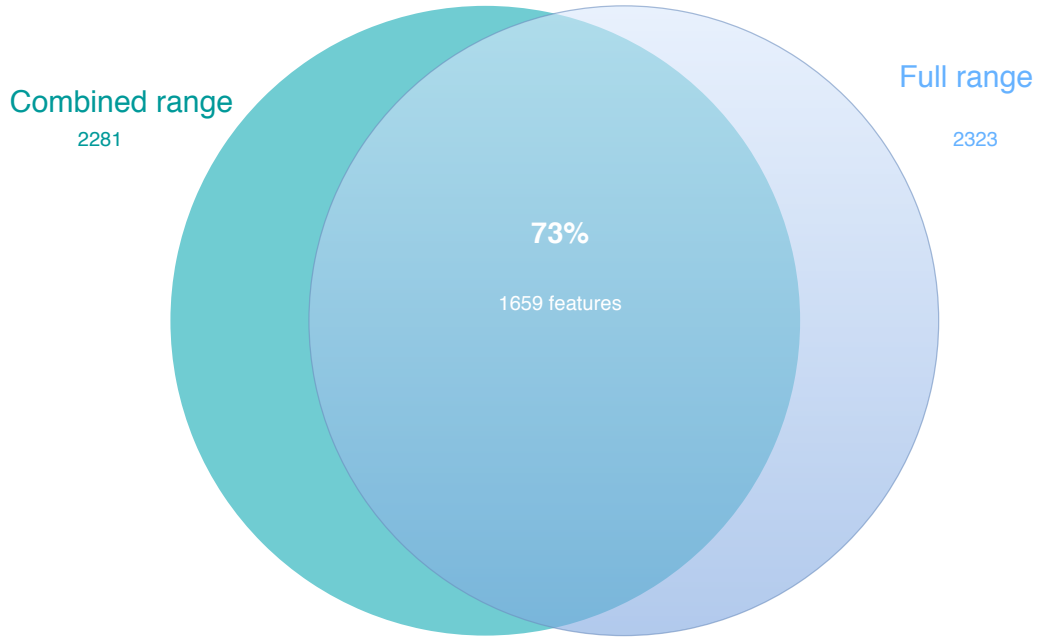


Figure 4.1 Comparison between "full" and "combined" mass ranges.
Venn diagram depicting total features found for "combined" and "full" *m/z* ranges as well as features common to found in both analysis types. A total of 1659 (73 %) features were common to both *m/z* ranges, with 2281 being found only in the combined range and 2323 in the full range.

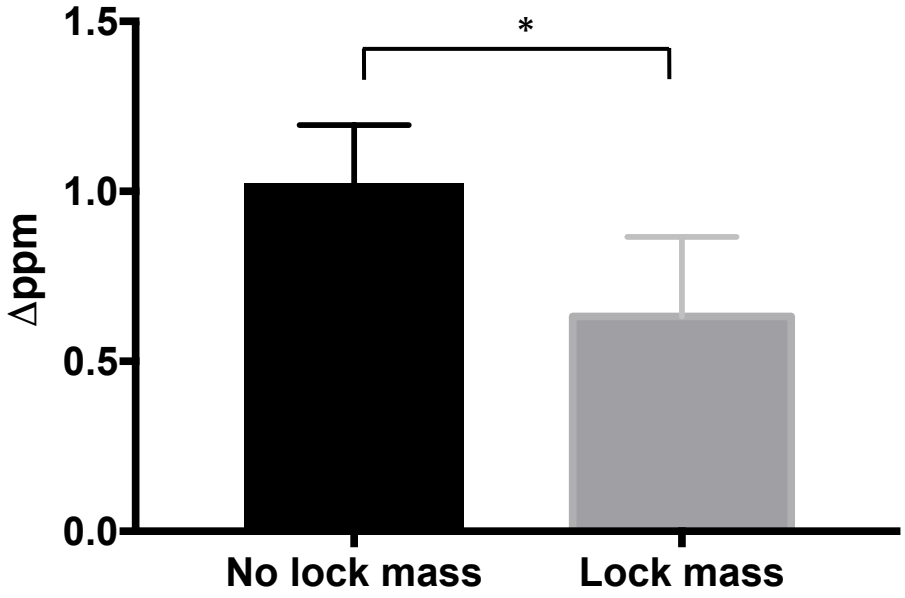


Figure 4.2 Increasing the mass accuracy of high-resolution LC-MS data using background ions.

Diisooctyl phthalate (*m/z* 391.2842) and sodium dodecyl sulphate (*m/z* 265.1479) were used as internal mass calibrants (lock-masses). Resulting in a significant improvement in mass accuracy. Bars represent the combination of all *m/z* values analysed in negative ion mode for samples analysed with and without lock mass with the respective SEM. Statistical significance determined using the Holm-Sidak method, with alpha=5.000%. Each row was analysed individually, without assuming a consistent SD, the graphs depict the mean and SEM. p-value, *<0.05, **<0.01.

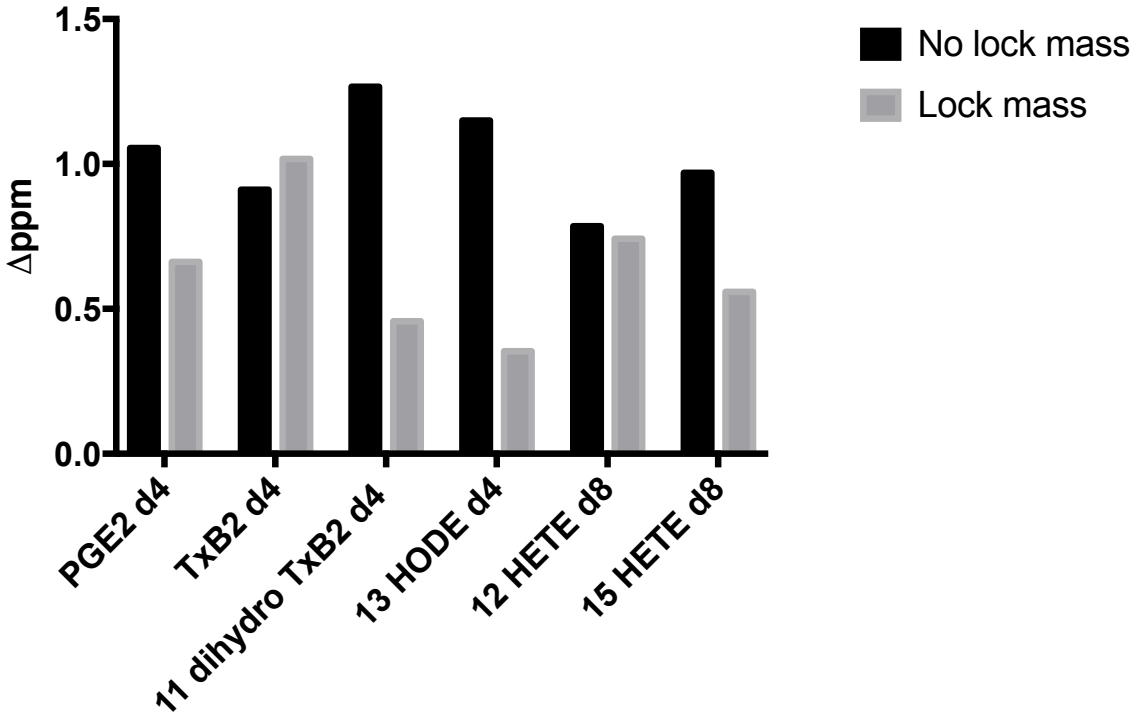


Figure 4.3 Increasing the mass accuracy of high-resolution LC-MS data using background ions.

Diisooctyl phthalate (*m/z* 391.2842) and sodium dodecyl sulfate (*m/z* 265.1479) were used as internal mass calibrants (lock-masses). Resulting in a significant improvement in mass accuracy. Bars represent the mass accuracy in negative ion mode, for samples analysed with and without lock mass, for different lipid standards. Individual deuterated internal standard Δppm values.

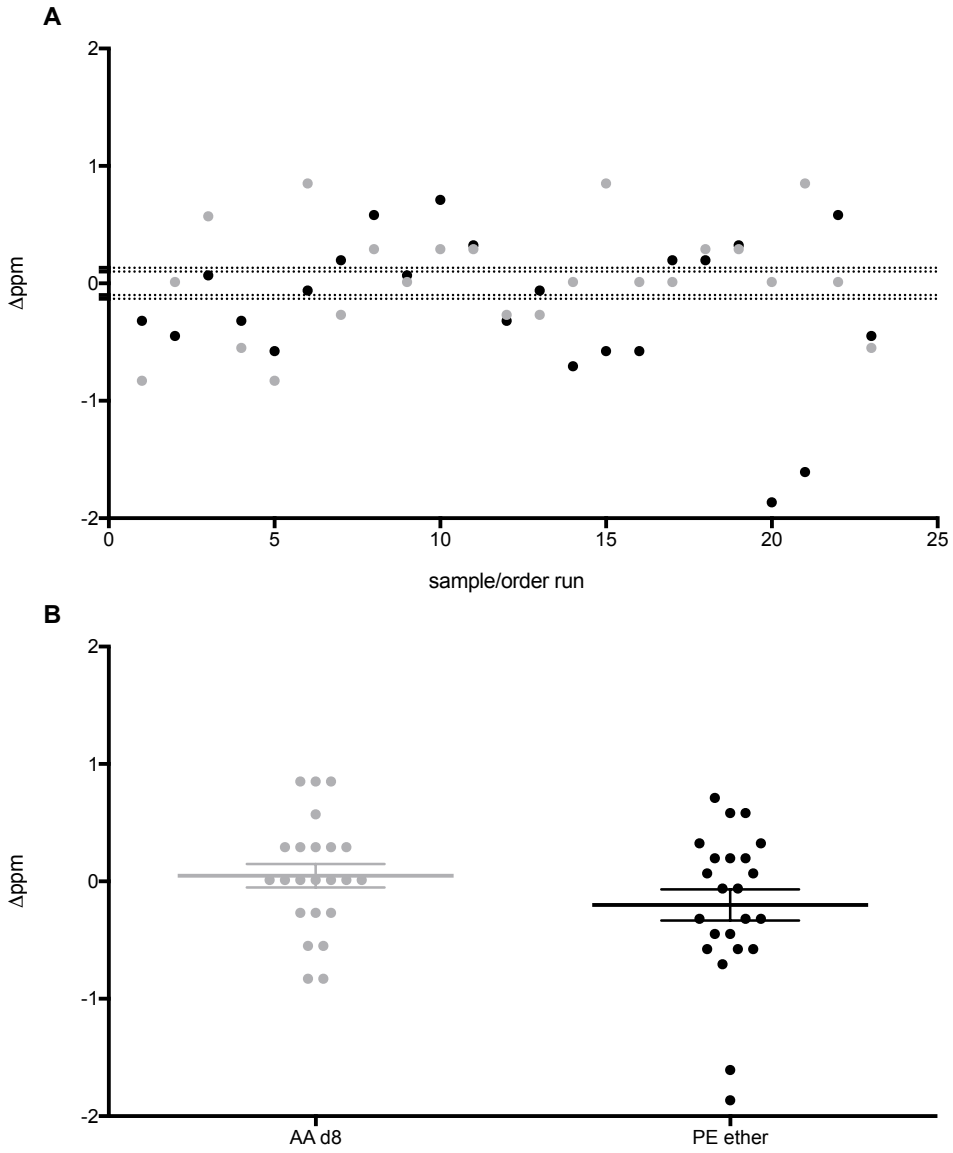


Figure 4.4 Increasing the mass accuracy and reproducibility of high-resolution LC-MS data using background ions.

A- scatter plot representation of mass accuracy variation for two deuterated internal standards with time (23 h). B- Mass reproducibility of samples depicted in panel A. Grey – Arachidonic acid d8, Black - PE ether.

4.2.2 Validation of LC-MS workflow

To compare the lipidomes of pResMϕ and RAW 264.7 cells, high-resolution LC- MS was used to maximize detection of lipids with close *m/z* values, and extensive chromatography was used to separate isobaric species and enable detection of low-abundance lipids.

pResMϕ were obtained from C57BL/6 mice aged between 10 – 13 weeks, through peritoneal lavage followed by Mϕ isolation using FACS sorting. RAW 264.7 cells were cultured as described in Chapter 2. Independent triplicates were collected for both cell types. Total lipids from these samples were extracted using a double extraction (Isopropanol/hexane/acetic acid followed by chloroform/methanol). The lipid extracts were then separated on an Accucore C18 column (150 x 2.1 mm, 2.6 μm) using an isocratic mobile phase A (H_2O:ACN 80:20 v/v) and a phase B (IPA: ACN, 70:30 v/v) at a flow rate of 0.425 mL/min over 60 min, with 4 mM ammonium acetate. High resolution (60,000, FWHM at 400 amu) full-scan spectra were acquired in both +ve and –ve ESI mode over 100 - 1800 *m/z* range on an Orbitrap (Elite). Chromatographic alignment, peak integration and isotope removal was done using XCMS. The dataset was then further processed using the software recently developed by the Cardiff Lipidomics group, LipidFinder[192]. Results were then queried against a lipid database, LIPID MAPS using the database search module integrated with LipidFinder with mass tolerance of 5 ppm.

4.2.3 Identification of Mϕ lipids using XCMS online

The processing of metabolomic data by XCMS Online is organized in three steps: data upload, parameter selection, and result interpretation[183]. The samples for both groups (pResMϕ and RAW cells) and extracted blanks (samples containing only PBS that went through the same extraction procedure as the cell samples) were uploaded through a specific Java applet, which allows the user to drag and drop the into the upload area of the program. During the file upload, parameters were selected matching the instrument setup used to analyse the samples (Methods). Once the files were uploaded and correct parameters were selected the 'job' was then submitted to the system.

XCMS Online uses nonlinear methods to compensate for retention time drifts between samples[193,194]. An overlay of all total ion chromatograms is shown before and after retention time correction, in addition to the retention time correction curves (Figure 4.5). After retention time correction, all TICs should be in alignment, allowing the differential analysis of lipids present in the 3 samples.

The "mirror plots"[183] in Figure 4.5D are part of the output from XCMS, show the dysregulated features, ions whose intensities are altered between the two Mϕ sample groups, according to the statistical threshold selected ($P < 0.1$ and $|FC| > 1.5$). Down-regulated features are represented as circles on the bottom of the plot, whilst up-regulated features are represented as circles on the top of the plot. The size of each circle corresponds to the fold change (up-regulated and down-regulated features are represented in green and red respectively); whereas the shade of the colour represents the p-value, with brighter circles having lower p-values. The retention time corrected TICs are also overlaid in the background of the figure.

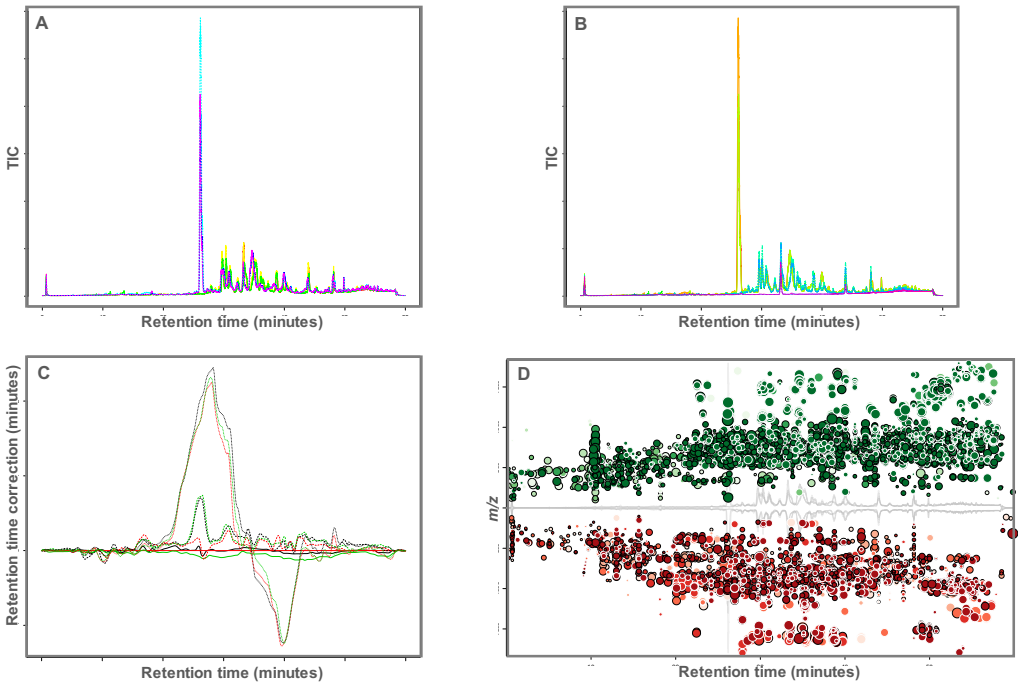


Figure 4.5 XCMS RT correction and "mirror plots".

Mϕ lipid extracts from RAW cells and pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. Overlay of all total ion chromatograms is shown before (A) and after retention time correction (B), retention time correction curves (C). "Mirror plots" show the dysregulated features, ions whose intensities are altered between the 2 Mϕ sample groups, according to the statistical threshold selected ($P < 0.1$ and $|FC| > 1.5$). Down-regulated features are represented as circles on the bottom of the plot, whilst up-regulated features are represented as circles on the bottom of the plot. The size of each circle corresponds to the (log) fold change (up-regulated and down-regulated features are represented in green and red respectively); shade of the colour represents the p-value, with brighter circles having lower p-values. The retention time corrected TIC are also overlaid in the background of the figure.

A table with details for all the identified lipid features is generated by XCMS (diffreport), as well as an overlay of the extracted ion chromatograms from each sample for all lipid features. This feature of XCMS is particularly important and it will be useful in further sections, as it was critical in the validation of the peak integration of the dataset.

XCMS provides putative identifications based on a METLIN search of the accurate mass, ordered by mass difference (ppm). However, METLIN is a computational database including over a million molecules ranging from lipids, steroids, plant & bacteria metabolites, small peptides, carbohydrates, exogenous drugs/metabolites, central carbon metabolites and toxicants; and whilst a large number of entries in a database is a positive factor, it can also be detrimental when there is no option to select a category such as lipids for a more concise search which often leads to incorrect putative identifications. Additionally, diffreport does not provide a reliable approximation of metabolites detected due to adducts, fragments, and artefacts. Indeed, depending on experimental conditions, more than 50 % of the features on a diffreport can be a result of fragments and artefacts[195].

Isotopic interferences arise when an isotope from one species occurs at the same nominal mass as the monoisotopic mass of another species[189]. This situation is very common in untargeted lipidomics approaches, which makes it essential to implement an accurate method of isotope correction (deisotoping)[196]. A final optional isotope removal step is done to identify and remove features, thus simplifying considerably the dataset. This step reduced the Mϕ lipidomics dataset size by *ca.* 33 % (Negative ion mode before deisotoping 3295, after 2216; Positive ion mode before deisotoping 10327, after 6849), Figure 4.6.

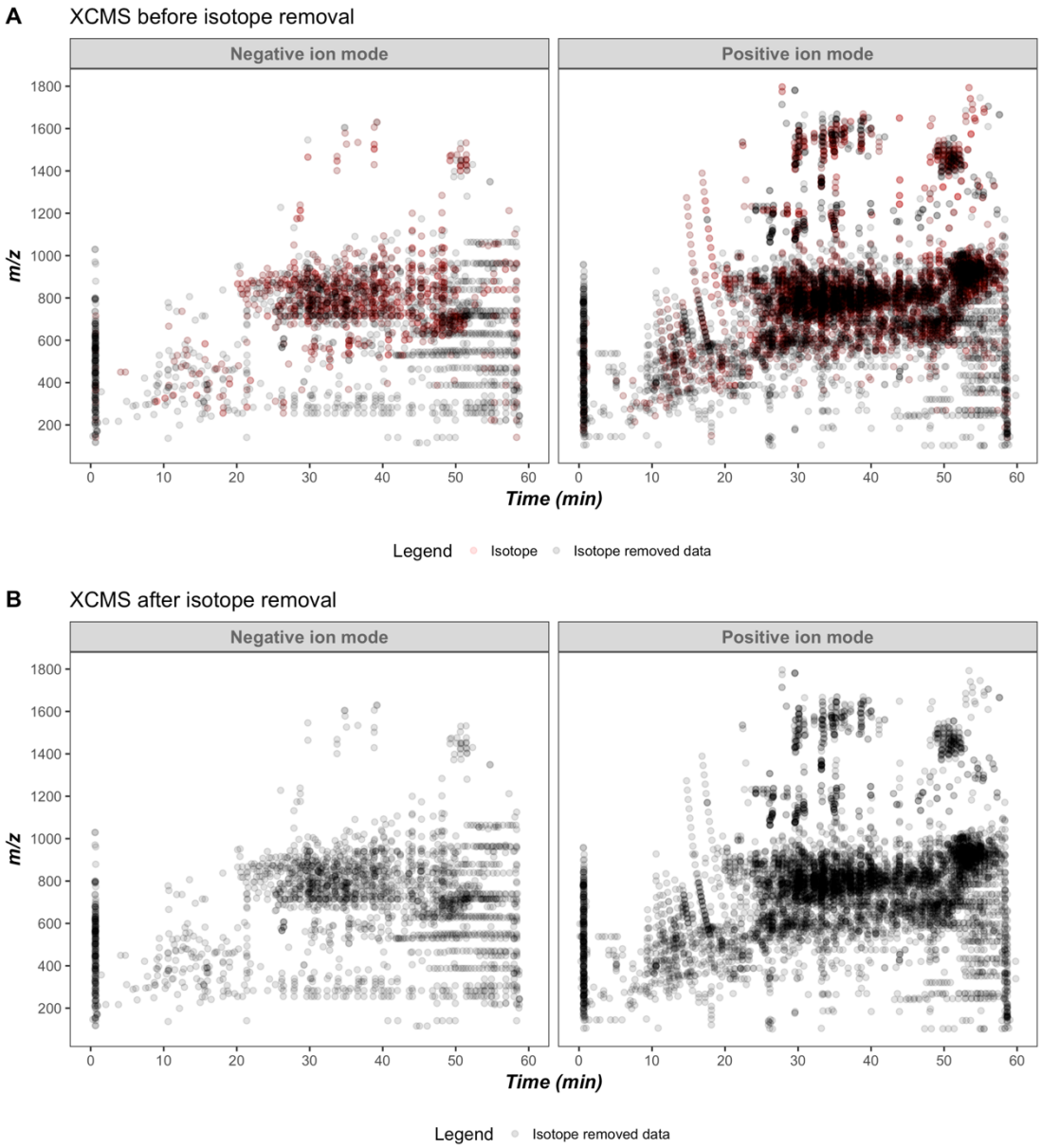


Figure 4.6 Isotope detection by XCMS on dataset containing ions present in both RAW cells and pResMϕ.

Mϕ lipid extracts from RAW cells and pResMϕ were analyzed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. Scatter diagrams showing elution of lipids and isotopes identified by XCMS, in either negative- or positive-ion mode. A – Mϕ removal of isotopes from the dataset processed with XCMS; B- Mϕ lipids after isotope removal. Red dots indicate *m/z* ions identified as isotopes removed during the isotope removal step; Black dots represent remaining *m/z* ions. Isotope removal reduced the Mϕ lipidomics dataset size by ca. 33% (Negative ion mode before deisotoping 3295, after 2216; Positive ion mode before deisotoping 10327, after 6849).

4.2.4 Computational workflow for discovery of lipids identifies significant differences in the lipidome of RAW 264.7 cells and pResMϕ

a) LipidFinder clean-up

Accurate and high-quality curation of lipidomic datasets is essential for cell biology studies and biomarker discovery. Nevertheless, a major challenge lies in removing artefacts otherwise mistakenly interpreted as real lipids from large LC- MS datasets, while retaining genuine ions in the dataset. Herein I used an approach specifically tailored for lipidomics, particularly discovery research, on the data obtained after XCMS processing. LipidFinder is an open-source Python workflow that screens the data for possible artefacts, and the outputs are assessed against online databases and categorized into LIPID MAPS categories[192].

A major issue with ESI, is the generation of ions arising from either common contaminating ions, in-source fragments of lipids, multiple noncovalent adducts, or solvent background[192]. The PeakFilter program is the main peak-finding algorithm in LipidFinder, and it includes contaminant/adduct removal and RT correction. The Mϕ lipid dataset was subjected to PeakFilter, resulting in a 39 % reduction of features, from 9065 to 5529 (Figure 4.7): 21 % solvent contaminants, 0.3 % mass contaminants and low intensity ions, 8.8 % adducts, 7.5 % outliers, 1.7 % RT exclusion ions.

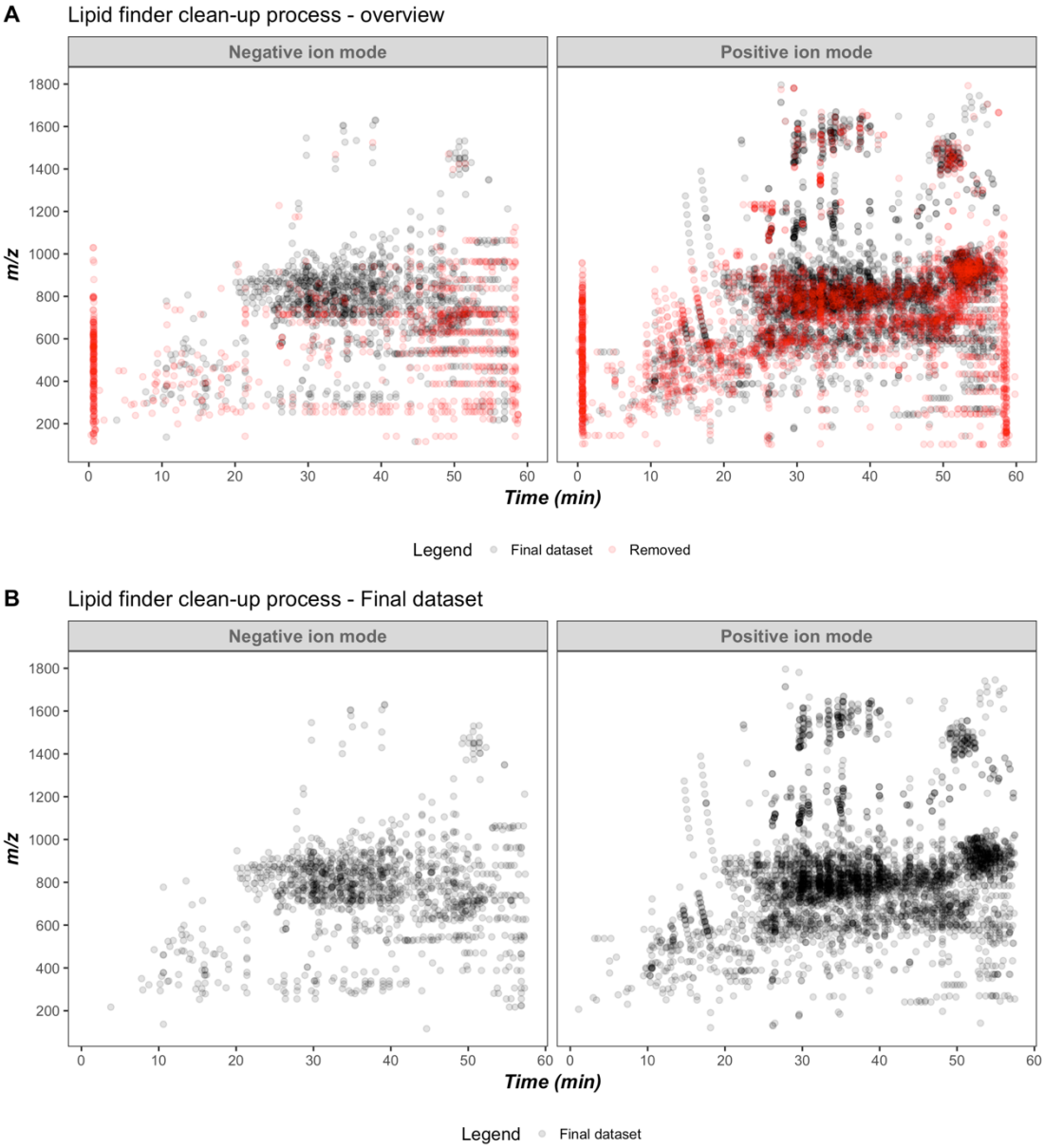


Figure 4.7 LipidFinder greatly reduces the number of features.

Mϕ lipid extracts from RAW cells and pResMϕ were analyzed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS followed by LipidFinder. Scatter diagrams showing elution of lipids identified by XCMS, in either negative- or positive-ion mode, highlighting the *m/z* ions removed by LipidFinder, resulting in a 39 % reduction of lipid features: 21 % solvent contaminants, 0.3 % mass contaminants and low intensity ions, 8.8 % adducts, 7.5 % outliers, 1.7 % RT exclusion ions. Red dots indicate *m/z* ions identified as artefacts and removed during PeakFilter; Black dots represent remaining *m/z* ions.

b) Putative identification of positively and negatively charged lipids by LipidFinder

The positive and negative ion mode data resulting from PeakFilter was then combined using the Amalgamator, an optional program used to combine data sets from positive and negative runs. The program found matches for 402 *m/z* ions resulting in a further reduction of the number of features by 3.7 %, from 5529 to 5324.

The MSSearch program in LipidFinder uses the list of *m/z* values resulting from the Amalgamator for putative identification based on accurate mass using the LIPID MAPS database[3] (Figure 4.8). However, there is still a large number of lipids missing from online databases and the results coming back as unknown is spread through the whole *m/z* range, accounting for about 24 % of the values searched. The putative identification process returned classifications for 76 % of *m/z* ions matching 5 of the 8 Lipid Maps categories, Fatty acyls (FA) 10 %; Glycerolipids (GL) 17 %; (GPL) 37 %; SL 12 %; Sterol lipids 0.36 %. Moreover, the lipids identified seem to overall match the *m/z* and retention times expected for the respective lipid categories.

4.2.5 Lipid levels and species vary greatly between pResMϕ and RAW cells

a) Global lipidomic differences between pResMϕ and RAW cells

Next, I examined the global differences between pResMϕ and RAW cells, analysing lipids whose levels are significantly different in both cell types (Figure 4.9). Of the 5265 features comprising the dataset, 69.4 % were significantly dysregulated. However, of the dysregulated features the number of significantly increased is surprisingly similar between the 2 samples. With RAW cells presenting a slightly higher percentage of upregulated, 55 % vs. the 45 % found in pResMϕ.

---

(1) [3] **COMP_DB**: a computationally generated database composed of over 30,000 bulk (isobaric) species covering 30 lipid classes which is customized for precursor ion searching.

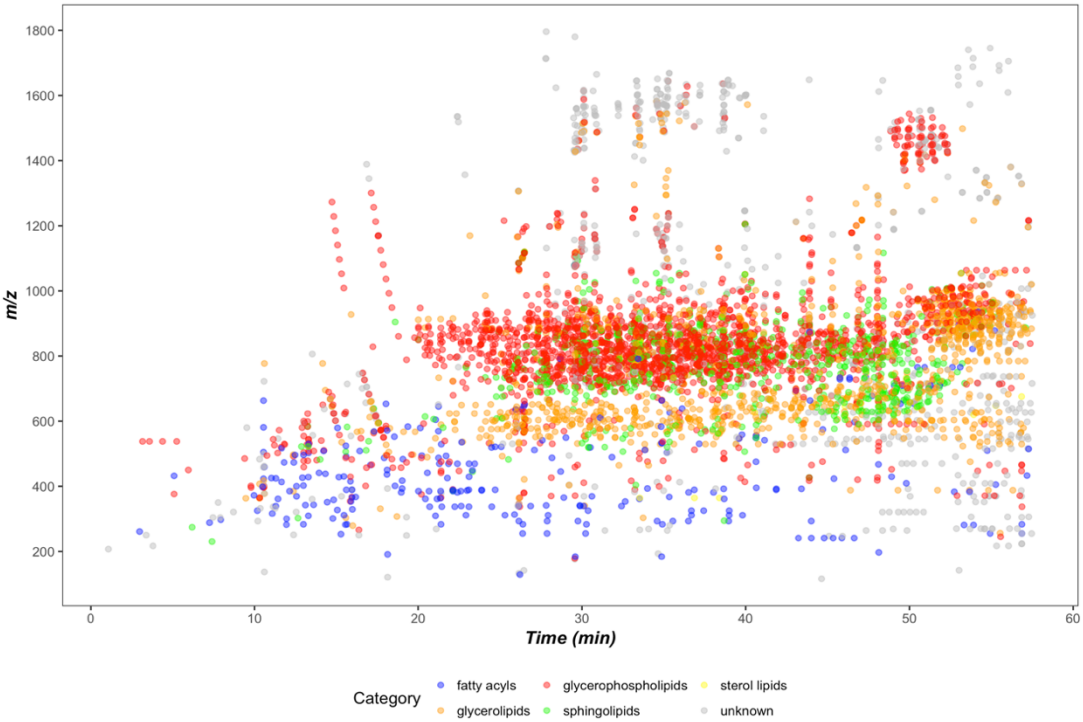


Figure 4.8 Putative identification of positively and negatively charged lipids by LipidFinder.

Mϕ lipid extracts from RAW cells and pResMϕ were analyzed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS followed by LipidFinder. Scatter diagram showing Mϕ lipids present in pResMϕ and RAW cells putatively identified using LipidFinder (MSSearch). Fatty acyls - 178 lipids (10 %); Glycerolipids – 1487 lipids (17 %); Glycerophospholipids -1974 lipids (37 %); Sphingolipids – 669 lipids (12 %); Sterol lipids – 17 lipids (0.36 %); Unknown – 940 lipids (24 %).

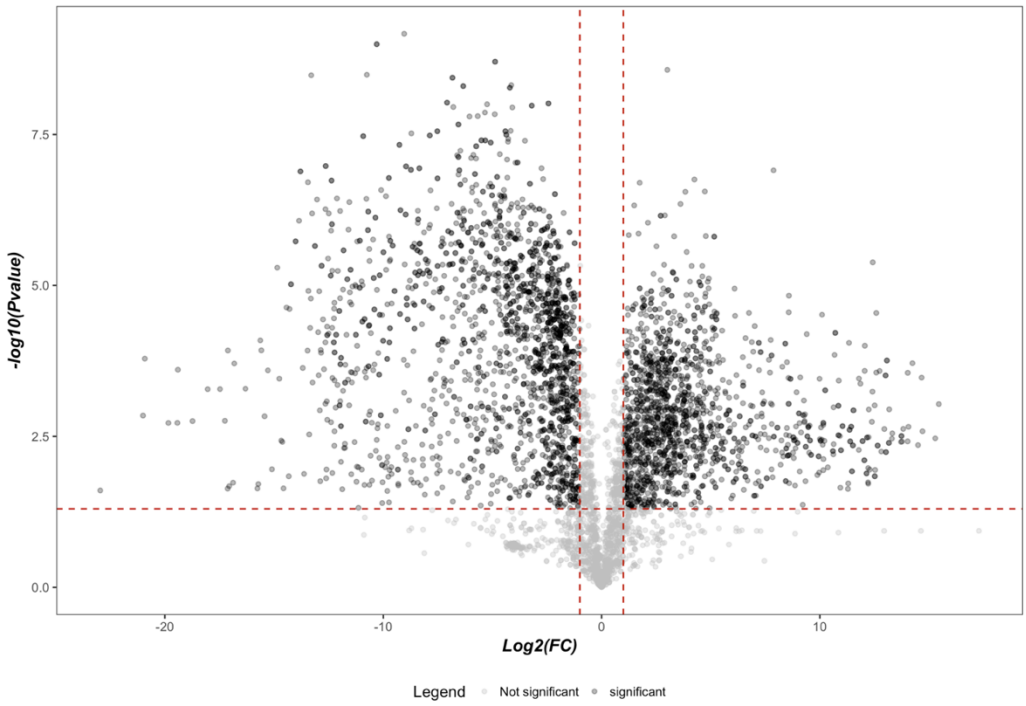


Figure 4.9 Lipid levels and species vary greatly between pResMϕ and RAW cells.

Mϕ lipid extracts from RAW cells and pResMϕ were analyzed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS followed by LipidFinder. Volcano plots representing fold-change and p-value of lipids in pResMϕ relative RAW cells. Significantly changed lipid levels are indicated by black (69.4 %), and unchanged levels by grey (30.6 %). Statistical significance was determined using the Holm-Sidak method, with alpha=5.0 %. Each row was analysed individually, without assuming a consistent SD. Differential expression was classified as significant when $P < 0.1$ and $|FC| > 1.5$ (depicted in black).

b) Lipid dysregulation by lipid category in Mϕ

Mapping the distribution of the upregulated features in pResMϕ and RAW cells by lipid category gave some insight into how these differ in the 2 cell types, leading me to take a closer look at how significant these differences are[4].

Volcano plots representing fold-change and p-value of lipids in pResMϕ relative RAW cells by lipid category, show that GL (73 %), GPL (68 %) and SL (66 %) are the categories with highest percentage of dysregulated lipids**.** Interestingly, FA (51 %) and GL (51 %) are the only categories where most of the significantly dysregulated lipids are upregulated in pResMϕ. However, even though not all the lipids were detected in both cell types, the percentage of lipids present only in one cell type remains below 10 % for most lipid categories (percentage of unique lipids calculated considering the significant *m/z* ions upregulated for pResMϕ and RAW cells). With the exception being the FA category, where 29 % of lipids upregulated in RAW cells are absent in pResMϕ.

Next, I will look into which lipid classes accounted for the highest percentages of molecules within the significantly changed lipid categories.

[4] Due to the low number of features identified as sterol lipids (17) it is not possible to assume significance of percentages.

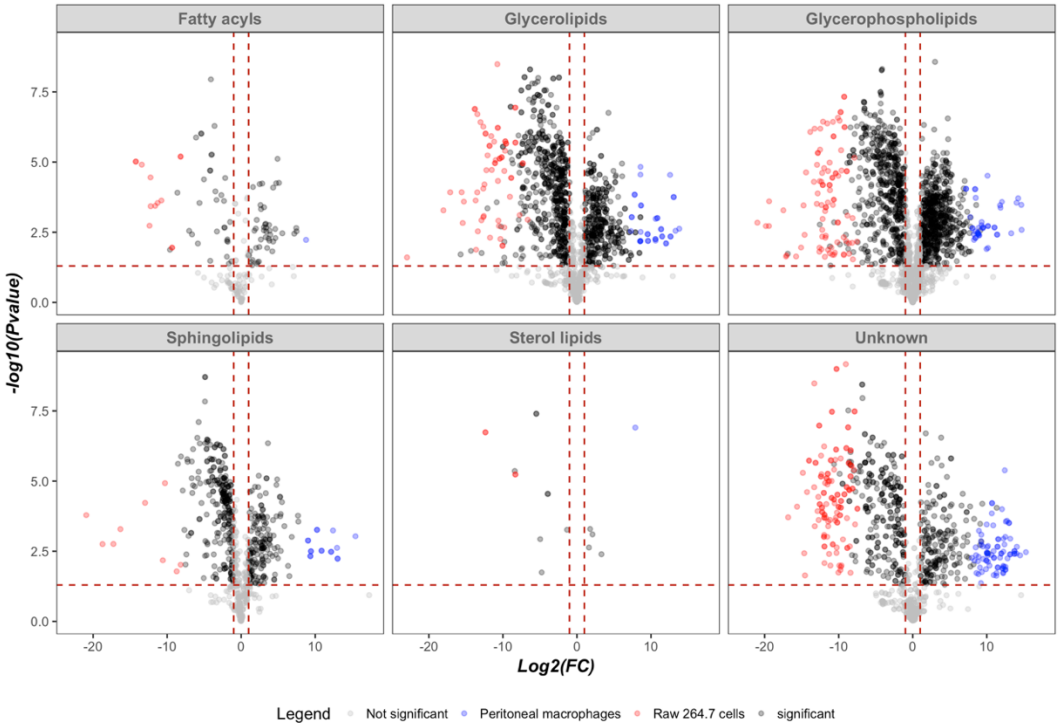


Figure 4.10 Volcano plot of significantly increased lipids in both pResMϕ and RAW 264.7 cells by lipid category.

Mϕ lipid extracts from RAW cells and pResMϕ were analyzed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS followed by LipidFinder. Volcano plots representing fold-change and p-value of lipids in pResMϕ relative RAW cells by lipid category. Significantly changed lipid levels are indicated by black, and unchanged levels by grey. Lipids found only in pResMϕ or RAW cells are coloured in blue and red, respectively. Statistical significance was determined using the Holm-Sidak method, with alpha=5.00%. Each row was analysed individually, without assuming a consistent SD. Differential expression was classified as significant when $P < 0.1$ and $|FC| > 1.5$ (depicted in black).

c) Significantly different lipid classes in Mϕ-like cells

Here I selected the significantly different lipids in pResMϕ and RAW 264.7 cells and found that the majority (44.4 %) were putatively identified as GPL, followed by GL (19.3 %) and SL (9.1 %) (Figure 4.11 inner ring). I then subdivided each lipid category into their respective lipid classes, which revealed that glycerophosphoethanolamines (PE) and glycerophosphocholines (PC) represented 20.7 % of all molecules, whilst triradylglycerols accounted for 8.3 %. Due to the structural similarities between PE and PC classes it is not possible to accurately distinguish them using this approach.

Next, I looked into the significantly changed lipid classes, searching for the ones that were particularly enriched in one cell type by plotting the ratio of the number of lipids significantly increased in RAW 264.7/pResMϕ (Figure 4.12). Thus, I found that only putatively identified bile acids and acidic GSL are increased in pResMϕ, whilst steroids, isoprenoids, diradylglycerols and Cer are mainly increased in RAW 264.7 cells.

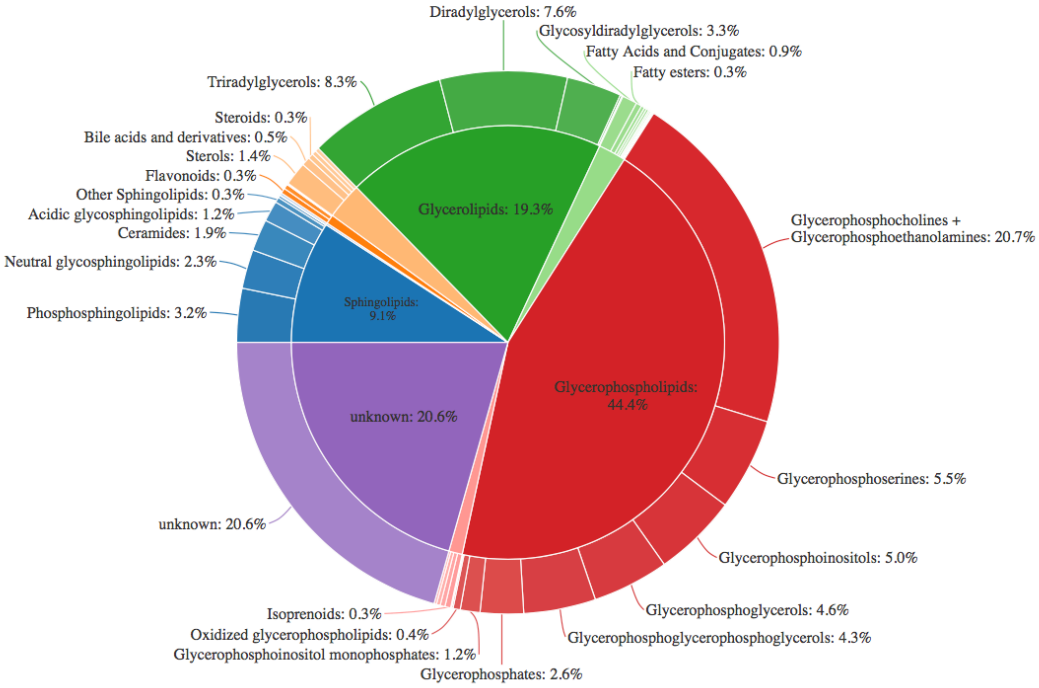


Figure 4.11 Global view of the lipidomic differences between of pResMϕ and RAW 264.7 cells.

The pie chart shows the percentage of the different lipid categories (inner ring) and lipid classes (outer ring) significantly altered between pResMϕ and RAW 264.7 cells, values below each category represents the percentage of lipids putatively identified for each category.

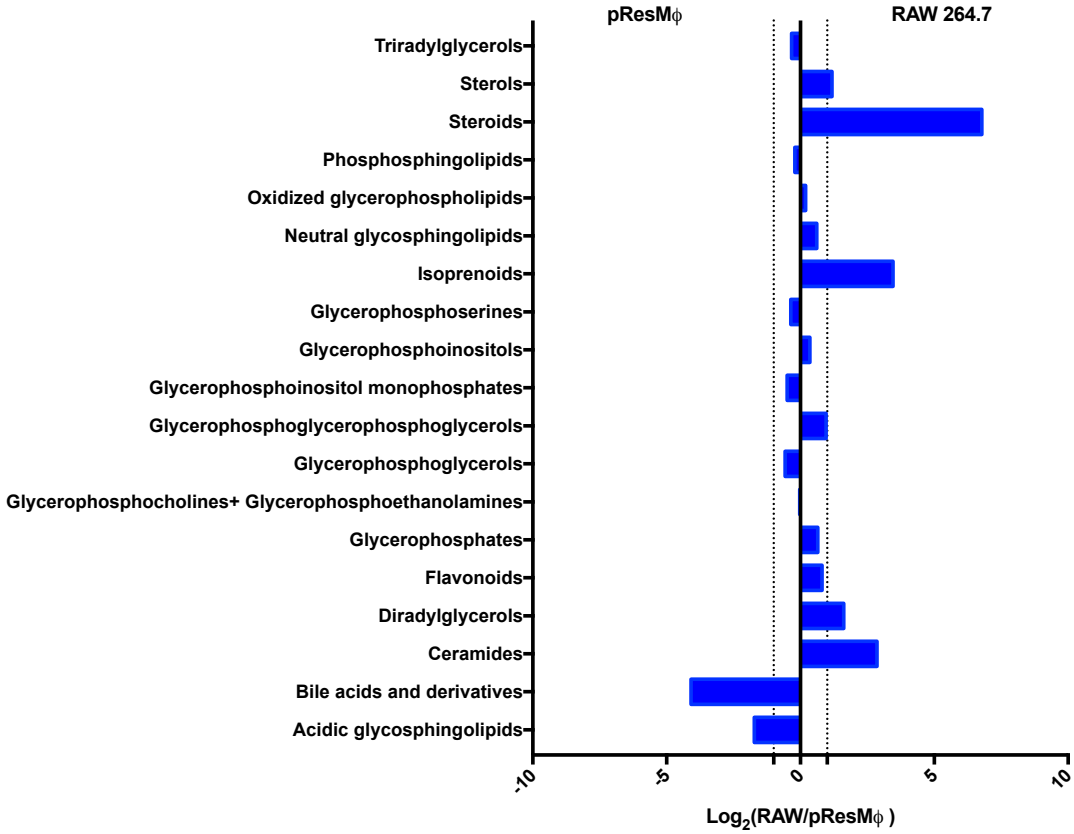


Figure 4.12 Lipid class enrichment in pResMϕ and RAW 264.7 cells.

The graph shows the enrichment of different lipid classes in pResMϕ and RAW 264.7 cells. The abscissa represents the Log_2 of the ratio of the number of lipids increased in RAW 264.7 and pResMϕ. Negative values- increased in pResMϕ; Positive values- increased in RAW 264.7.

4.3 Discussion

In this Chapter the aim was to optimise the high-resolution untargeted LC-MS lipidomics workflow, previously implemented in platelet studies, for the analysis of Mϕ extracts. Due to the different nature and low abundance of the samples analysed, it was necessary to optimise the different steps of the workflow, such as LC-MS conditions and data analysis.

In order to efficiently profile the lipidome of Mϕ-like cells, it is necessary to reach the lowest mass accuracy possible[189]. To this end, I employed the lock mass feature, which successfully improved the Δppm to 2 ppm. However, one of the drawbacks of using the untargeted approach is that the methods are often associated with relatively long runtimes, as is the case of the LC-MS runs in this thesis (60 min). Herein I combined 2 mass ranges, previously used for other studies, into a single mass range (from 100 – 1800 *m/z*), described in section 4.2.1a), thus reducing the number of chromatographic runs and the amount of sample necessary to be cut in half.

One of the main challenges for lipidomics, in particular for high-resolution untargeted LC-MS-based approaches, resides in the computational and bioinformatics difficulties of handling the large amount of data obtained from such experiments. Specifically, high-resolution untargeted LC-MS data collection requires substantial efforts in chromatographic alignment and statistical appraisal of variations in signal intensity. These fluctuations can have different origins, such as biological variations, sample handling and analytical accuracy[104]. Subsequently, several replicates are usually required for accurate determination of lipid changes in complex matrixes by high-resolution untargeted LC-MS.

Since their discovery Mϕ have one of the most studied immune cell types because of their central role in immune processes and disease. However, the technical expertise required to process LC-MS-based untargeted metabolomic data has caused lipidomics to lag behind metabolomics because it was mainly undertaken by specialists, such as biologists or chemists with primary interests in the biology or chemistry of lipids, who did not have an analytical background.

Nevertheless, lipids make up about a third of the metabolome, so the field is now catching up. One of the drawbacks of high-resolution LC-MS analysis, as the method I performed in this Chapter, is the number of artefacts and contaminants contained in the raw data. In order to 'clean-up' my data-set and minimise the number of false- positive lipid features I took advantage of the XCMS Online peak filtering and deisotoping capabilities and LipidFinder's clean-up and MSSearch capabilities, which lead to a reduction 39 % reduction of lipid features. Thus, providing me with a more comprehensive and biologically relevant dataset.

The results presented herein hint at the fact that distinct Mϕ phenotypes are significantly different metabolically, 69.4% of lipid features were found to be significantly dysregulated. The categories with highest percentage of dysregulated lipids were GL (73 %), GPL (68 %) and SL (66 %)**.** Interestingly, FA (51 %) and GL (51 %) are the only categories where most of the significantly dysregulated lipids are upregulated in pResMϕ. A closer analysis of which lipid putatively identified classes were enriched in the different Mϕ subsets revealed that: (i) bile acids and derivates are significantly increased in pResMϕ. These lipids act as signalling molecules, affecting lipopolysaccharide (LPS)-induced cytokine expression in Mϕ, and are able to reprogram pro-inflammatory Mϕ to an anti-inflammatory phenotype[197,198]. (ii) the steroid class is significantly increased in RAW 264.7 when compared to pResMϕ. Steroids are synthesized from cholesterol and have long been known to stimulate phagocytic activity in Mϕ, and are able to downregulate pro-inflammatory cytokine production[199,200]. These two findings suggest that RAW 264.7 might have a more pro-inflammatory phenotype, while pResMϕ may have an increased phagocytic activity and a more anti-inflammatory/pro-resolving phenotype.

Overall the work developed in this Chapter was extremely important for the development of an adequate workflow for the lipidomics analysis of pResMϕ. It enlightened which features, processes, and conditions are useful in a differential analysis of two Mϕ samples. The conditions optimised herein will be carried on to the next Chapter where this workflow will be used to characterise the lipidomic differences in *Gata6*-WT and -KO^{mye} pResMϕ.

Chapter 5 Global lipidomic characterisation of *Gata6*-WT and -KO^{mye} pResMϕ

5.1 Introduction

To address the question of whether myeloid deficiency of GATA6 is associated with significant changes in lipid metabolism, at a transcriptional and molecular level, I conducted experiments to establish a connection between *Gata6* expression and alterations in lipid metabolism-related pathways/networks (Chapter 3). Thus, I confirmed that the knockdown of *Gata6* expression leads to significant changes at a transcriptional level, with emphasis on SL related causal networks.

In this Chapter I will extend my transcriptional analysis, by characterising the global lipidome of *Gata6*-WT and -KO^{mye} pResMϕ. This will allow me to determine (in conjunction with Chapters 4) the fundamental changes both in lipid levels and potential signalling deficits associated with the myeloid deficiency of GATA6, as well as which lipid categories are the most influenced by *Gata6*.

I will begin by isolating pResMϕ from *Gata6*-WT and -KO^{mye}. Then I will perform an untargeted lipidomic characterisation, using the workflow developed in Chapters 4, which uses high-resolution mass spectrometry coupled to ultra-high pressure liquid chromatography (UPLC-MS). The data will then be analysed using computational methodologies combined with bioinformatics tools in order to remove artefact signals, search databases, and provide putative classification for lipids found. Last, multi- and univariate data analysis will be employed to determine statistical significance.

5.1.1 Aim

- To characterise the global lipidome of *Gata6*-KO^{mye} and –WT pResMϕ.
- To determine the effect of GATA6 deficiency on lipid molecular species in peritoneal Mϕ.

5.2 Results

5.2.1 GATA6 regulates lipid profiles in pResMϕ

Lipid extracts were analysed by UPLC-MS (as described in Chapter 2). Briefly, peritoneal lavages collected from *Gata6*-WT and -KO^{mye} female mice were phenotyped and pooled[5] (details in Materials and Methods, Chapter 2). pResMϕ were isolated by FACS (pooled to provide 2 x 10^6 cells/replicate), centrifuged, and cell pellets stored at -80°C, until ready for extraction. When all three biological replicates were collected, lipids were extracted using a double extraction method (Hexane/Isopropanol; Bligh and Dyer). I then used a lipidomics platform, consisting of a Nexera UPLC coupled with an Orbitrap Elite, to profile the lipid extracts from *Gata6*-WT and -KO^{mye} (Figure 5.1A).

The raw data obtained from the UPLC-MS analysis, was processed using chromatographic alignment, peak integration and isotope removal by XCMS. The dataset was then manually inspected, and incorrectly integrated features were removed. A table with details for all identified lipid features is generated by XCMS (diffreport), as well as an overlay of the extracted ion chromatograms from each sample for all lipid features. This feature of XCMS is particularly important and it was critical in the validation of the peak integration of the dataset (First and second stages Figure 5.1B).

The positive and negative ion mode data resulting from XCMS was submitted to a program that combines data sets from positive and negative runs (LipidFinder, described in the previous Chapter). This program found matches for 4.4 % of ions, causing a further reduction of the number of features. The final number of features was 14988.

[5] Due to the low recovery of pResMϕ in GATA6 deficient mice, it was necessary to combine peritoneal lavages from ~15 mice. In order to obtain 2 x 10^6 cells/replicate.

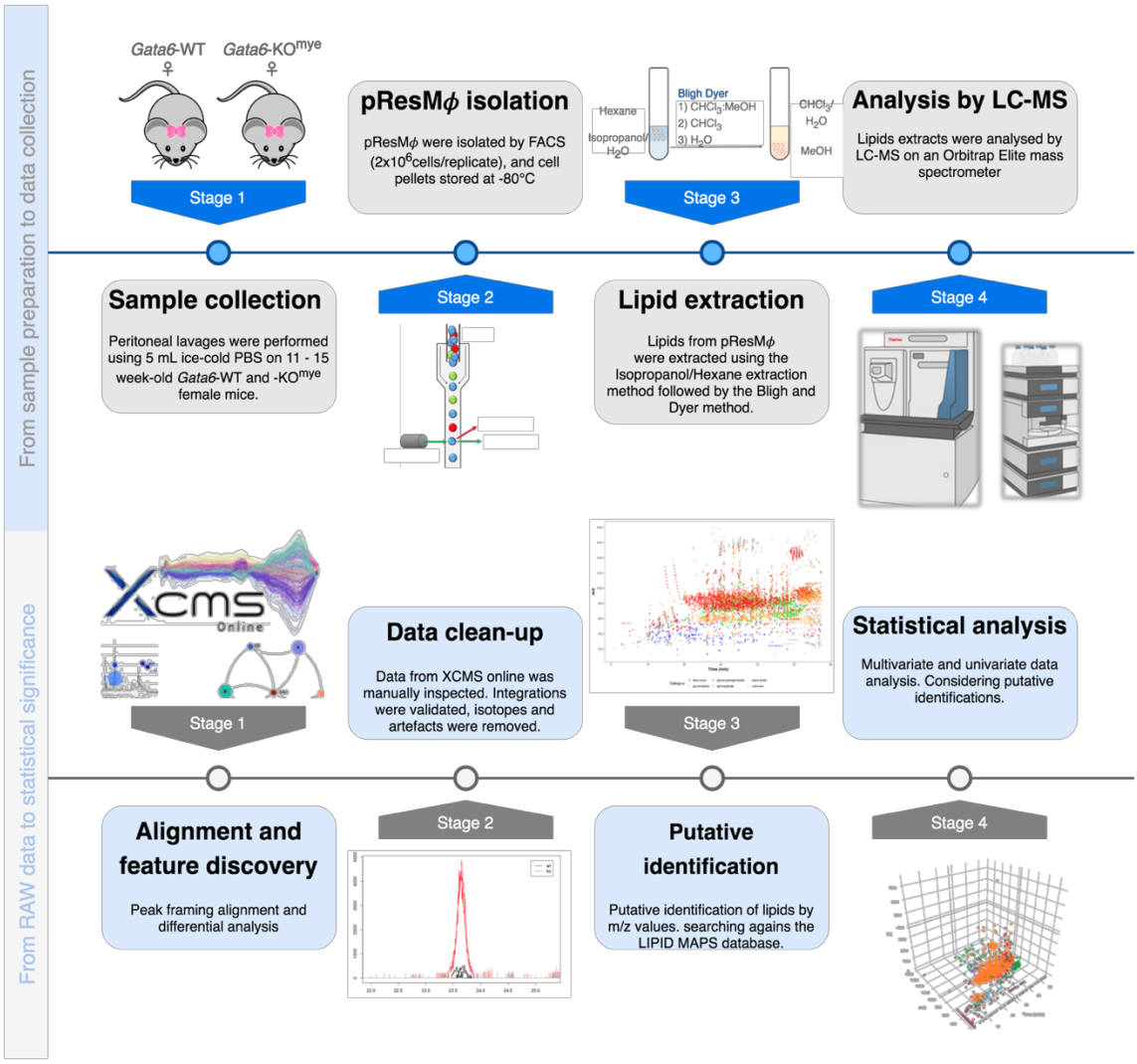


Figure 5.1 From sample collection to statistical significance workflow.

The figure depicts the different stages of the untargeted lipidomics workflow described herein. Top panel- describes the steps involved from sample preparation to data collection, including sample collection, cell isolation, lipid extraction and analysis by LC-MS, which is then followed by the steps described in the bottom panel, From Raw data analysis to statistical significance. This comprises alignment and feature discovery, data clean-up, putative identification and statistical analysis.

a) Power analysis

Power analysis is performed to determine the minimum sample size used to detect an effect size of interest, with power defined as the probability of detecting an effect, given that the effect is really there. For example, if a study comparing two groups (*e.g*. WT vs KO) has power of 0.8, then 80 % of the time, the analysis was performed, it would deliver a statistically significant difference between the two groups.

This calculation can be affected by three main factors:

- Effect size which is usually defined as the difference of the two group means divided by the pooled standards deviation.
- Degree of confidence for statistical significance (α).
- Sample size, more samples will increase power.

Ideally this analysis would have been performed using pilot data[201]. However, due to the limitations encountered when working with primary cells, in particular *Gata6*-KO^{mye} and –WT pResMϕ, where the yield of cells is ~50 % lower than with WT animals, I was limited to 3 biological replicates per group. Thus, I used the full dataset for the calculations. Nevertheless, it is important to investigate the power of my experimental design, as it will help me calculate the magnitude of the effect of interest in the population (effect size).

As metabolomics evolves, and becomes more widely used, careful consideration of methods, to ensure proper experimental design are essential for an accurate and robust identification of metabolites. Traditional power analysis tools have become unsuitable for high-throughput metabolomics data. The nature of this data means that it is highly dimensional and often correlated. Moreover, the number of metabolites identified greatly outnumbers the sample size[201]. This data will be used as a screen, that will serve to select the most significantly altered lipid categories. The changes observed will then be validated in the next Chapter using a targeted approach.

In order to assess the power of the *Gata6*-WT and -KO^{mye} dataset, I used both a traditional analysis tool as well as a modified method of power analysis, that uses the average power of all metabolites, and corrects it for multiple testing using false discovery rate (FDR) instead of raw p-values.

First, I chose 10 lipid features at random, and calculated the sample size (for an α of 0.05 and power of 80 %) for each lipid. Resulting in a sample size larger than 3 samples/group for 7 of the selected features, Table 5.1. Extrapolating these results to the entire dataset would suggest that a sample size of 3 will result in 30 % of lipid features being significantly different.

Next, I used the entire untargeted lipidomics dataset and processed it using the power analysis module in Metaboanalyst[202]. The output consisted of four diagnostic plots (Figure 5.2), where it is possible to evaluate whether or not the calculations were successful. The analysis is based on the assumptions that the effect is present in the data, and that test statistic follows a normal or near normal (Students't t-test) distribution

The FDR selected, or significance criterion selected was α= 0.1, which resulted in a predicted power of 34 % for a sample size of 3 (Figure 5.2A), suggesting that my experimental design is underpowered, but it will still be able to detect significant differences between the 2 groups. In order to achieve 80 % power, the sample size would have to be 16, which would correspond to culling approximately 480 animals (15 animals/sample). The majority of p-values is close to zero (Figure 5.2B and D), indicating a strong effect, whilst the QQ plot confirms that the data follows a normal distribution (the data points follow the straight line) (Figure 5.2C).

Table 5.1 Power calculations for *Gata6*-WT and -KO^{mye} untargeted lipidomics data.

Peaks (mz/RT)	Mean WT	Mean KO	Stdev WT	Stdev KO	Sample size
788.544/29.14	8.95E+07	1.56E+08	1.64E+07	1.02E+07	2
431.379/24.33	6.26E+03	4.50E+04	5.85E+03	1.80E+04	5
932.5859/19.75	5.85E+04	7.30E+04	2.89E+04	3.95E+03	30
789.5471/21.55	1.03E+05	6.55E+05	1.63E+04	4.11E+04	2
964.6236/48.02	2.28E+04	4.22E+04	9.51E+03	8.75E+03	5
405.2654/16.11	4.11E+03	1.95E+04	6.52E+02	2.22E+04	30
854.4816/36.99	4.37E+05	3.18E+05	8.32E+04	2.12E+05	30
1625.132/35.23	4.76E+05	5.86E+05	8.44E+04	1.65E+05	30
789.686/52.81	1.52E+05	2.06E+05	5.01E+04	9.55E+04	30
680.5441/27.27	2.56E+06	6.87E+04	1.66E+04	1.57E+04	2

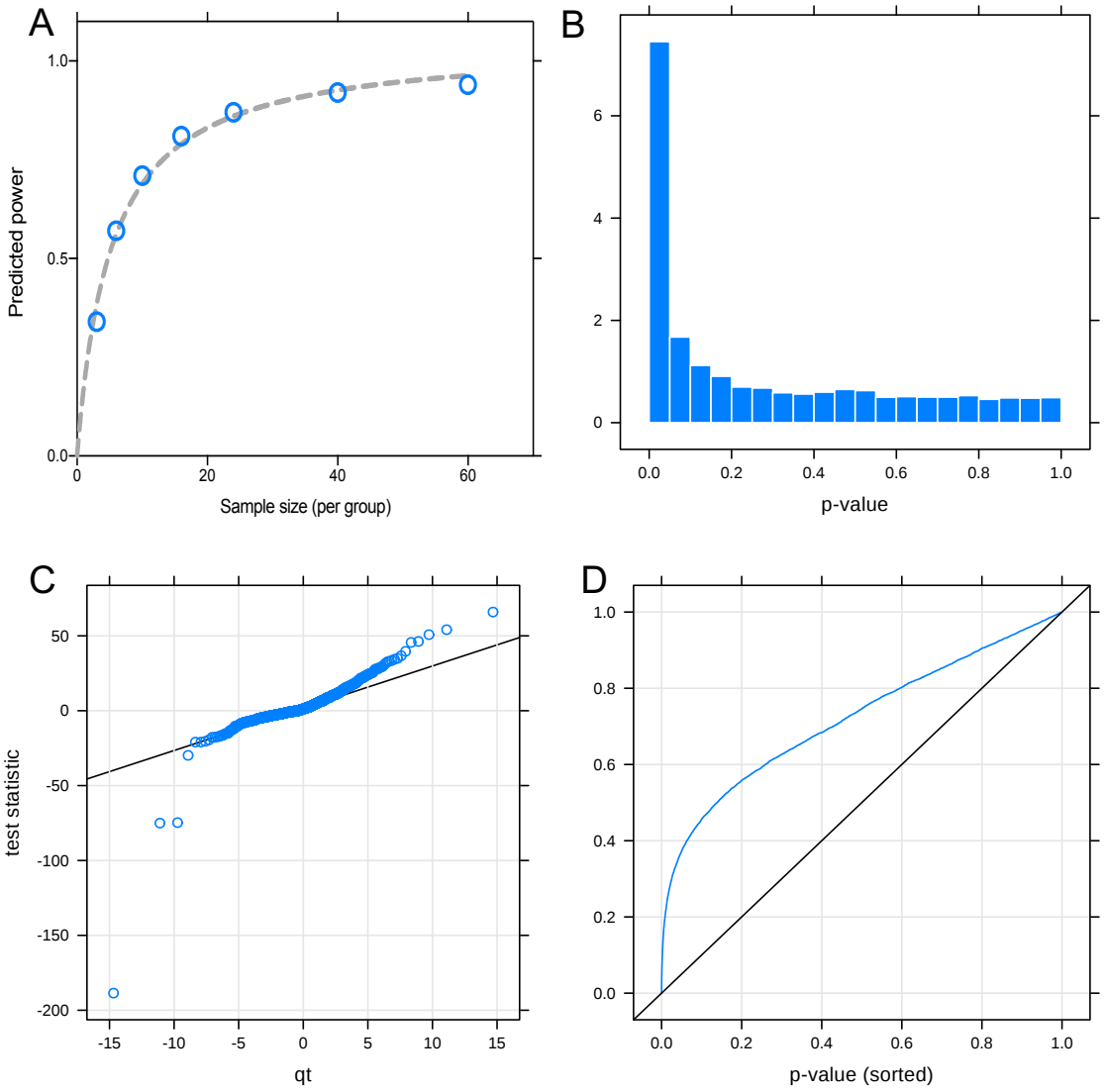


Figure 5.2 Power statistics of untargeted lipidomics data comparing *Gata6*-WT and -KO^{mye}.

Mϕ lipid extracts from *Gata6*-WT and -KOmye pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. Various plots overviewing the test-statistics and p-values calculated from the data to evaluate if it follows a normal distribution. A- Plot of the predicted power curve. B- shows the distribution of the p-values, where the majority of p-values hover around 0. C & D- Quantile-Quantile plots (QQ plots), the test-statistics and the p-values follow a standard normal distribution (the data points follow the straight line). For the qq-plot, the p-values are sorted against their ranks.

b) Statistical and Machine Learning Data Analysis

i) Univariate analysis methods

Once the dataset was processed by XCMS (as described in the previous Chapter), I performed a statistical analysis to identify lipid features that comprehend the bulk of the variance within the dataset. The selection of an suitable statistical method is dependent on the sample size and distribution or the ability to evaluate the normality of the distribution.

Univariate analysis methods are the most common methods used for exploratory data analysis. Herein I did a Fold Change (FC) analysis, t-tests, Mirror plot, and volcano plot, which is a combination of the first two methods. The univariate analyses provide a preliminary overview about features that are potentially significant in discriminating the conditions under study.

Fold Change Analysis

For paired FC analysis, the algorithm used first counts the total number of pairs with fold changes that are consistently above/below the specified FC threshold (|FC| > 1.5) for each variable. A variable is reported as significant if this number is above a given count threshold (> 75 % of pairs/variable) (Figure 5.3).

Of note, the purpose of FC is to compare absolute value changes between two group means. Therefore, the result is plotted in $\log_2$ scale, so that same fold change (up/down regulated) will have the same distance to the zero baseline.

Figure 5.3A shows the 3062 features identified as significant by fold change analysis, which represents 20.3 % of the dataset. The putative identification column shows the results obtained using the LIPID MAPS database. The LIPID MAPS database enabled a lipid category to be assigned, with a putative match generated, where possible, using an in-house database. The results show that 15 of the 20 features with the highest FC are GPL (5), SL (2), FA (2), and GL (1). For the remaining 5 lipid features it was not possible to find a match in the database.

T-tests

The second univariate method used was a Student t-test, or unequal variances t-test, which is a is a two-sample location test. The purpose of t-test analysis is to test whether two distributions come from the same population[203].

Figure 5.3B shows the 1897 features identified by t-tests (P-value < 0.1) (P-values were corrected for multiple-testing using Benjamini-Hochberg). The details of the 20 features with the lowest P-value are shown Table 5.3. The method used for putative identification of lipid features was as described in the previous section. The results show that 17 of the 20 features with the smallest P-values are GPL (11), SL (2), FA (2), and ST (2). For the remaining 3 lipid features it was not possible to find a match in the database. Of note, none of the features selected are present in both the FC and the t-test analysis (top 20 features).

In the next section I will provide further insight about which features are potentially significant in discriminating *Gata6*-WT and -KO$^{\mathbf{mye}}$ pResMϕ using a "mirror plot" analysis.

Mirror plots

The "mirror plots"[183] in Figure 5.4, show the dysregulated features, ions whose intensities are significantly altered between *Gata6*-WT and -KO pResMϕ, according to the statistical threshold selected (p-value < 0.1 and |FC| > 1.5). Down-regulated features are represented as red circles on the bottom of the plot, whilst up-regulated features are represented as green circles on the top of the plot. The size of each circle corresponds to the fold change (up-regulated and down-regulated features are represented in green and red respectively); whereas the shade of the colour represents the p-value, with brighter circles having lower p-values. The retention time corrected TIC are overlaid in the background of the figure. This analysis revealed an even spread of lipid features throughout the 5 and 50 min RT range, with more features presenting increases in intensity (upper part of the plot).

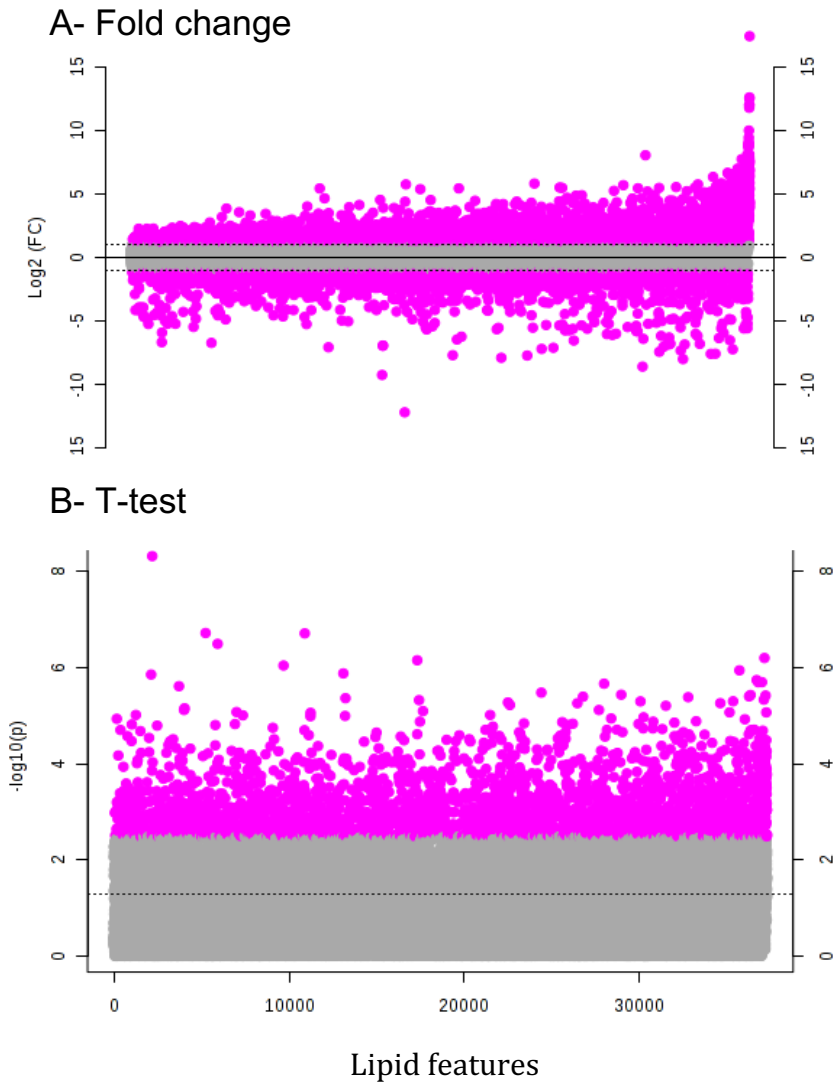


Figure 5.3 Fold-change and t-test analysis displays evidence of lipid regulation by GATA6.

Mϕ lipid extracts from *Gata6*-WT and -KO^{mye} pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. Figure A shows important features selected by fold-change analysis with threshold 2. The pink circles represent features above the threshold. Note the values are on log scale, so that both up-regulated and down-regulated features can be plotted in a symmetrical way. Figure B shows important features selected by t-tests with threshold 0.1. The pink circles represent features above the threshold. Note the p values are transformed by -log10 so that the more significant features (with smaller p values) are plotted higher on the graph.

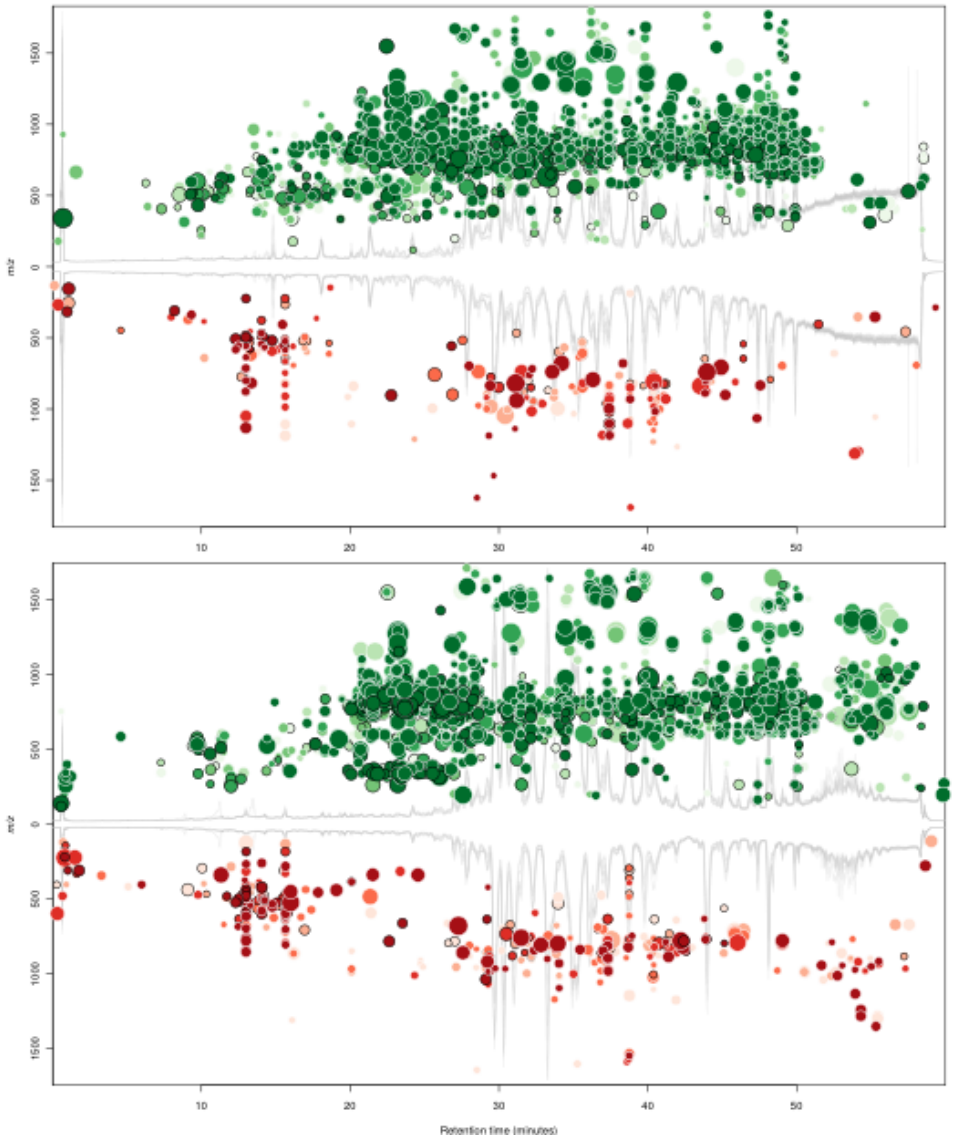


Figure 5.4 "Mirror plots" show evidence of lipid regulation by GATA6.

Mϕ lipid extracts from *Gata6*-WT and -KO pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. "Mirror plots" show the dysregulated features, ions whose intensities are altered between the 2 Mϕ sample groups, according to the statistical threshold selected (p-value < 0.1 and |FC| > 1.5). Down-regulated features are represented as circles on the bottom of the plot, whilst up-regulated features are represented as circles on the bottom of the plot. The size of each circle corresponds to the (log) FC (up-regulated and down-regulated features are represented in green and red respectively); shade of the colour represents the p-value, with brighter circles having lower p-values. The retention time corrected TIC are also overlaid in the background of the figure. (A)Positive ion mode; (B) Negative ion mode.

Volcano Plot

Volcano plots use a combination of FC and t-test analysis, where important features are selected based on the fold change threshold (x) 2 and t-tests threshold (y) 0.1. The further its position away from the (0,0), the more significant the feature is. Figure 5.5 shows the important features identified by volcano plot, 8.4 % of the whole dataset, of which 91.2 % were increased in GATA6 deficient cells. Table 4 shows the details of the 20 most significantly altered features. These features match exactly with the ones identified by t-test.

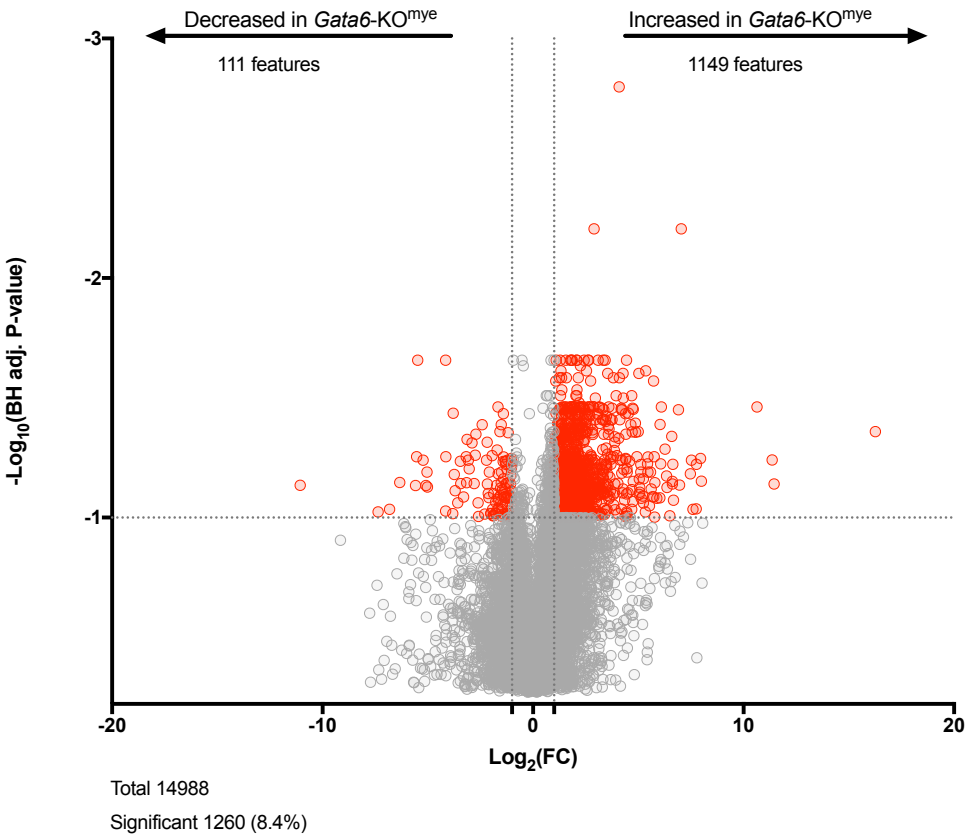


Figure 5.5 Lipid levels vary significantly between *Gata6*-WT and -KO pResMϕ.

Mϕ lipid extracts from *Gata6*-WT and -KO pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. Volcano plots representing fold-change and p-value of lipids in *Gata6*-WT relative to -KO pResMϕ. Significantly changed lipid levels are indicated by red (8.4 %) (of which 91.2 % were increased in GATA6 deficient cells), and unchanged levels by grey (87.4 %). Statistical significance was determined using the student t-test, with alpha= 10.00 %, P-values were corrected for multiple-testing using Benjamini-Hochberg. Each row was analysed individually, without assuming a consistent SD. Differential expression was classified as significant when P< 0.1 and |FC| > 1.5 (depicted in red). Note both fold changes and p-values are log transformed. The further its position away from the (0,0), the more significant the feature is.

i) Multivariate analysis methods:

Data reduction approaches such as principle component analysis (PCA) or supervised approaches, including orthogonal partial least squares discriminant analysis (OPLS-DA), can help to reduce complex data to a manageable form, identify associations within datasets and identify key metabolites contributing to the variance[204].

Multivariate data analysis (MDA) by PCA provided an overview of the contribution of GATA6 to the different lipid classes and further complements the results from section 5.2.1c). MDA methods are typically used for group differentiation in lipidomics, using either supervised or unsupervised approaches. PCA reduces primary variables (peaks) to the latent variables, which are called principal components, due to the decreasing model dimensionality[205].

Data scaling

The first step before performing an MDA is to scale the data, in particular in cases where the variables being compared have different response units. For LC-MS data, all of the responses correspond to peak areas and hence have the same units, so scaling is not absolutely essential, but is still useful. If the original data is not scaled in any way the variables with the largest response will tend to dominate the PCA analysis. For example, a very large variable which is approximately constant for all samples may dominate PC1, or as is the case with this dataset, peaks measured in positive ion mode tend to be more intense than those detected in negative mode, thus skewing the MDA results. For this reason, I used the Pareto scaling method, where the data is first mean centered and then divided by the square root of the standard deviation for the variable. The net effect is that larger variables receive more importance than with autoscaling, but less than with mean centering alone, Figure 5.6.

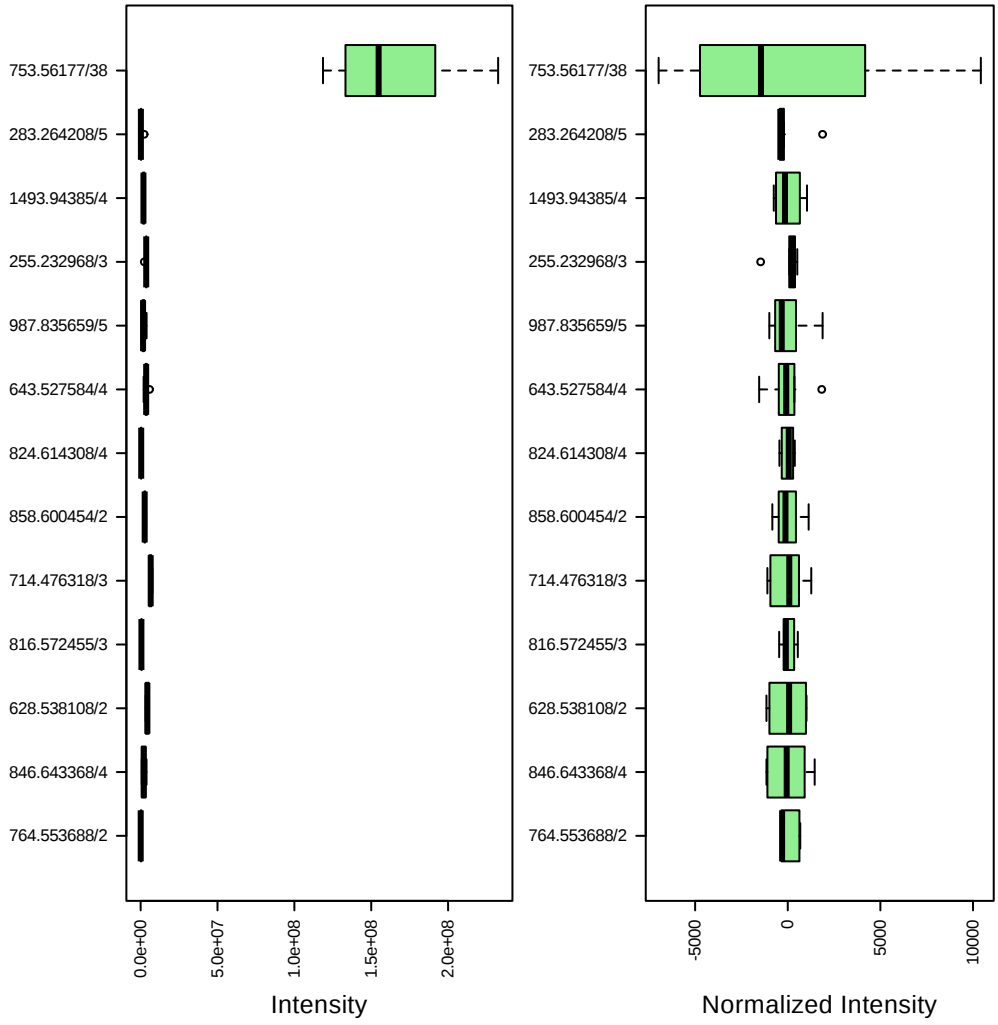


Figure 5.6 Untargeted lipidomics data of *Gata6*-WT and -KO^{mye} pResMϕ scaled using the Pareto method.

Mϕ lipid extracts from *Gata6*-WT and -KO^{mye} pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. The boxplots show 13 features due to space limit. The density plots are based on all samples.

ii) Principal Component Analysis (PCA)

PCA is an unsupervised method that aims to find the directions that best explain the variance in a data set (X) without considering class labels (Y). The data are summarized into less variables called scores (one variable per sample) which are weighted average of the original variables. The weighting profiles are called loadings. Herein, the PCA analysis was performed using Metaboanalyst, and the calculations are based on singular value decomposition.

Figure 5.7 shows pairwise score plots providing an overview of the various separation patterns among the most significant principal components (PC) (from 1 to 5), where PC1 explains 62.1% of the variance between *Gata6*-KO^{mye} and -WT. The PCA scree plot of the variance explained by PC1-PC5 showed that the other significant components are given by PC2-PC5, with percentages of contribution of 19 %, 8.2 %, 6.1 %, and 4.7 %, respectively. By removing the PCs that contribute little to the variance (PC3-5), I can project the entire dataset to a lower dimensional space but retain most of the information (Figure 5.8), and allow for the clustering of the two sample groups (Figure 5.9).

Because the PCA is a transformation of the old coordinate system (peaks) into a new coordinate structure (PC), it can be assessed how much each of the old coordinates (peaks) contribute to each of the new ones (PCs). The higher the loading (further away from the cluster) of a particular peak onto a PC, the more impact they have on the variance between *Gata6*-KO^{mye} and -WT. Similarly, to what was found during the univariate analysis, the majority of the features were identified as GPL (13) and SL (6).

In the next section I will look into how *Gata6*-KO^{mye} affects lipid levels in different categories, which will allow me to direct my studies to a particular lipid category.

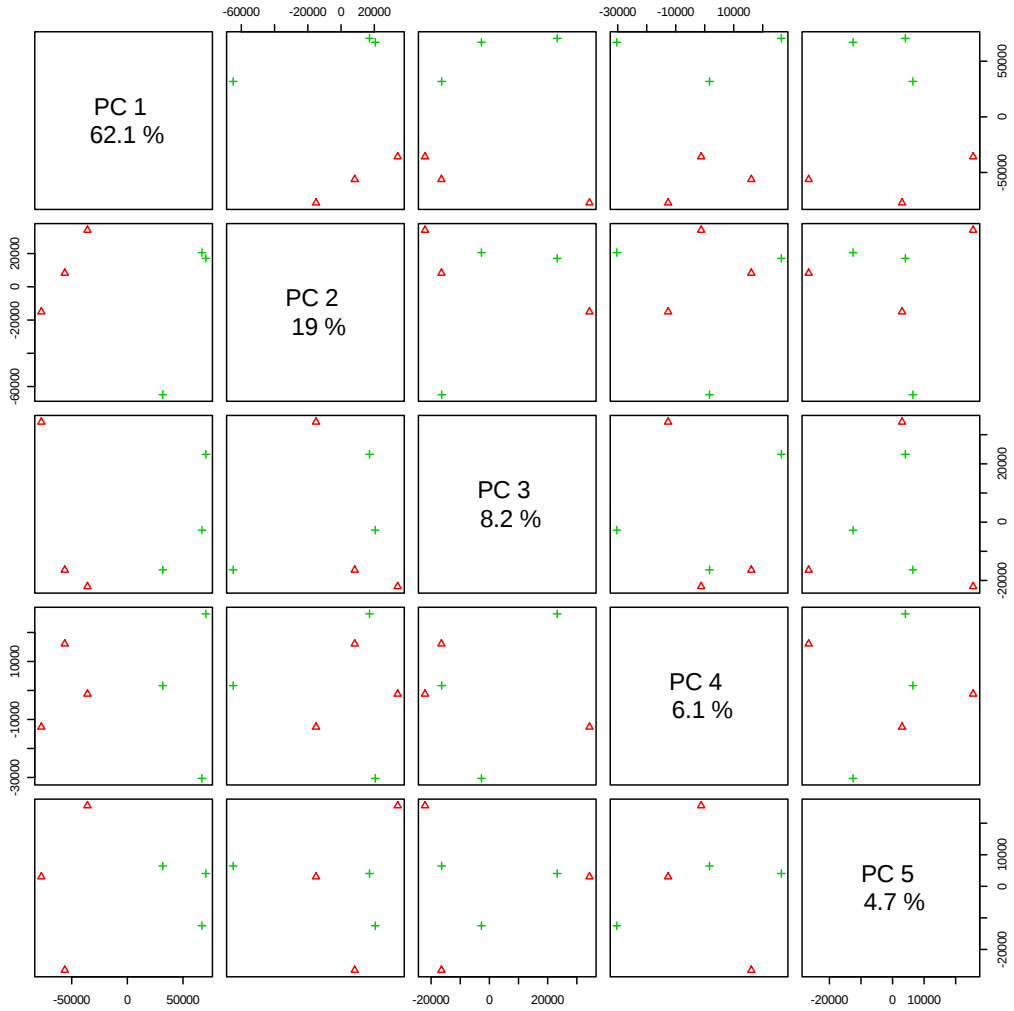


Figure 5.7 PC 1 explains 62.1 % of the variability between *Gata6*-KO^{mye} and -WT.

Mϕ lipid extracts from *Gata6*-WT and -KO pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. The figure shows pairwise score plots between the selected PCs. The explained variance of each PC is shown in the corresponding diagonal cell. Green and red symbols represent individual samples, *Gata6*-WT and -KO^{mye} respectively.

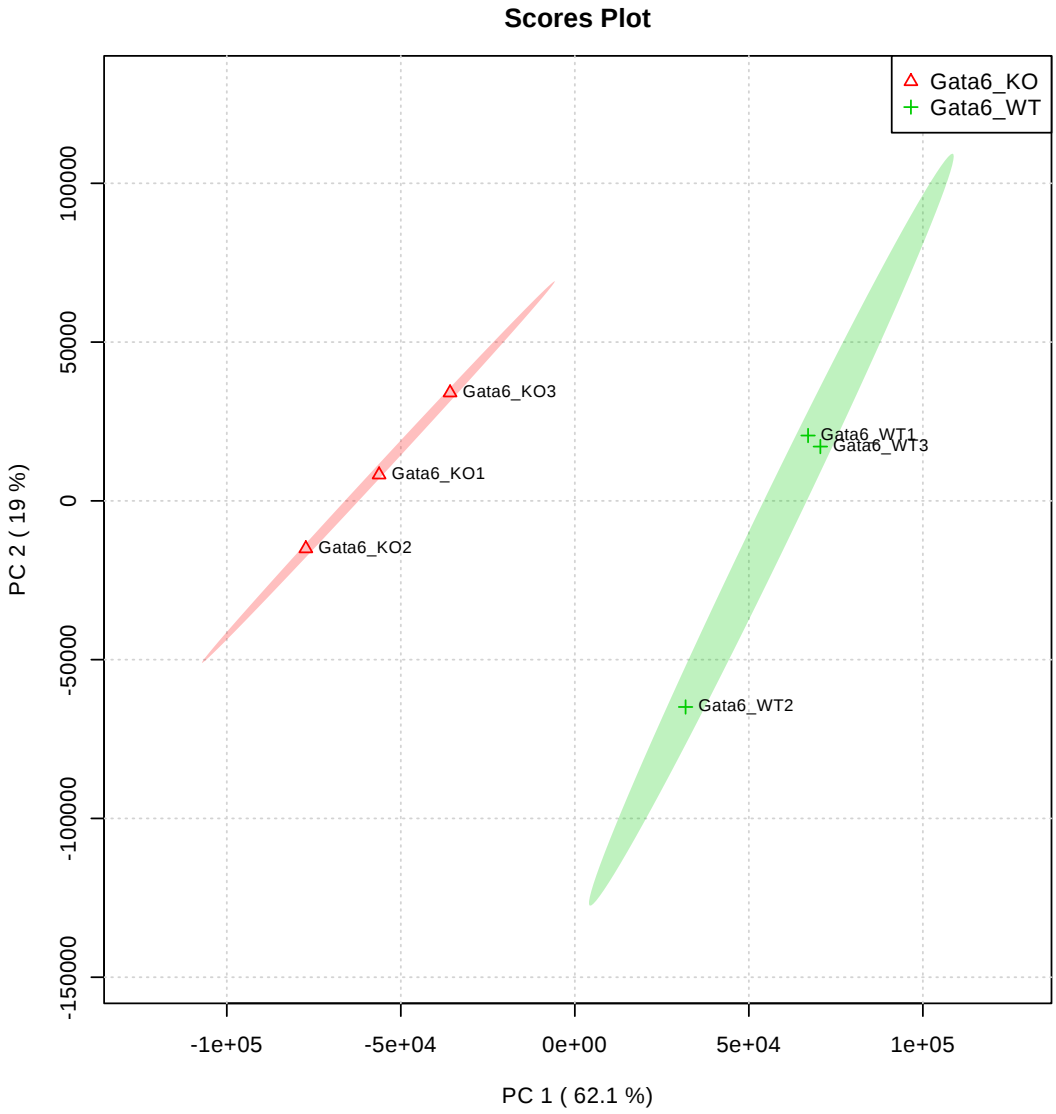


Figure 5.8 PCA analysis shows a clear separation between *Gata6*-KO^{mye} and -WT.

Mϕ lipid extracts from *Gata6*-WT and -KO pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. The figure shows pairwise score plots between the PC1 and PC2. Green and red symbols represent individual samples, *Gata6*-WT and -KO^{mye} respectively. And shaded areas represent the 95% confidence intervals corresponding to each group. The explained variances are shown in brackets.

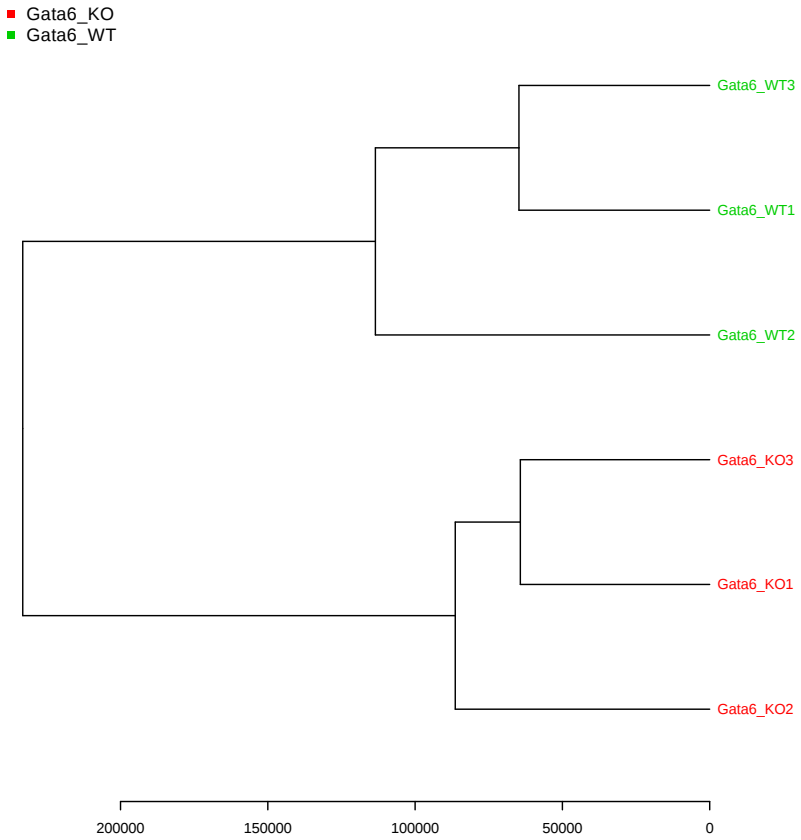


Figure 5.9 Unsupervised MDA clusters *Gata6*-WT and -KO^{mye} samples into their respective groups.

Untargeted lipidomics data was submitted to PCA analysis and clustering. Result shown as dendrogram (distance measure using euclidean, and clustering algorithm using ward.D).

c) GATA6 deficiency leads to significant changes in lipid levels in various lipid categories in pResMϕ

As previously mentioned (Chapter 4), the MSSearch program in LipidFinder uses the list of *m/z* values resulting from the Amalgamator for putative identification based on accurate mass using the classification system from the LIPID MAPS database (Figure 5.10). Herein I used the MSSearch program to putatively identify all the features in the dataset, Figure 5.11. However, there is still a large number of lipids missing from online databases and as such the number of results coming back as unknown is higher than what would be desirable, and spread through the whole *m/z* range, accounting for about 24.1 % of the values searched. The putative identification process returned classifications for 75.9 % of *m/z* ions matching the 8 Lipid Maps categories. The categories with the most matches were FA 13 %, GL 9 %, GPL 32 %; SL 10 % (Figure 5.11). Moreover, the lipids identified seem to overall match the *m/z* and retention times expected for the respective lipid categories. However, this analysis gives no information regarding the molecular abundance of the different categories, or their differences in *Gata6*-WT and KO^{mye} cells. These will be analysed in further detail in the following sections.

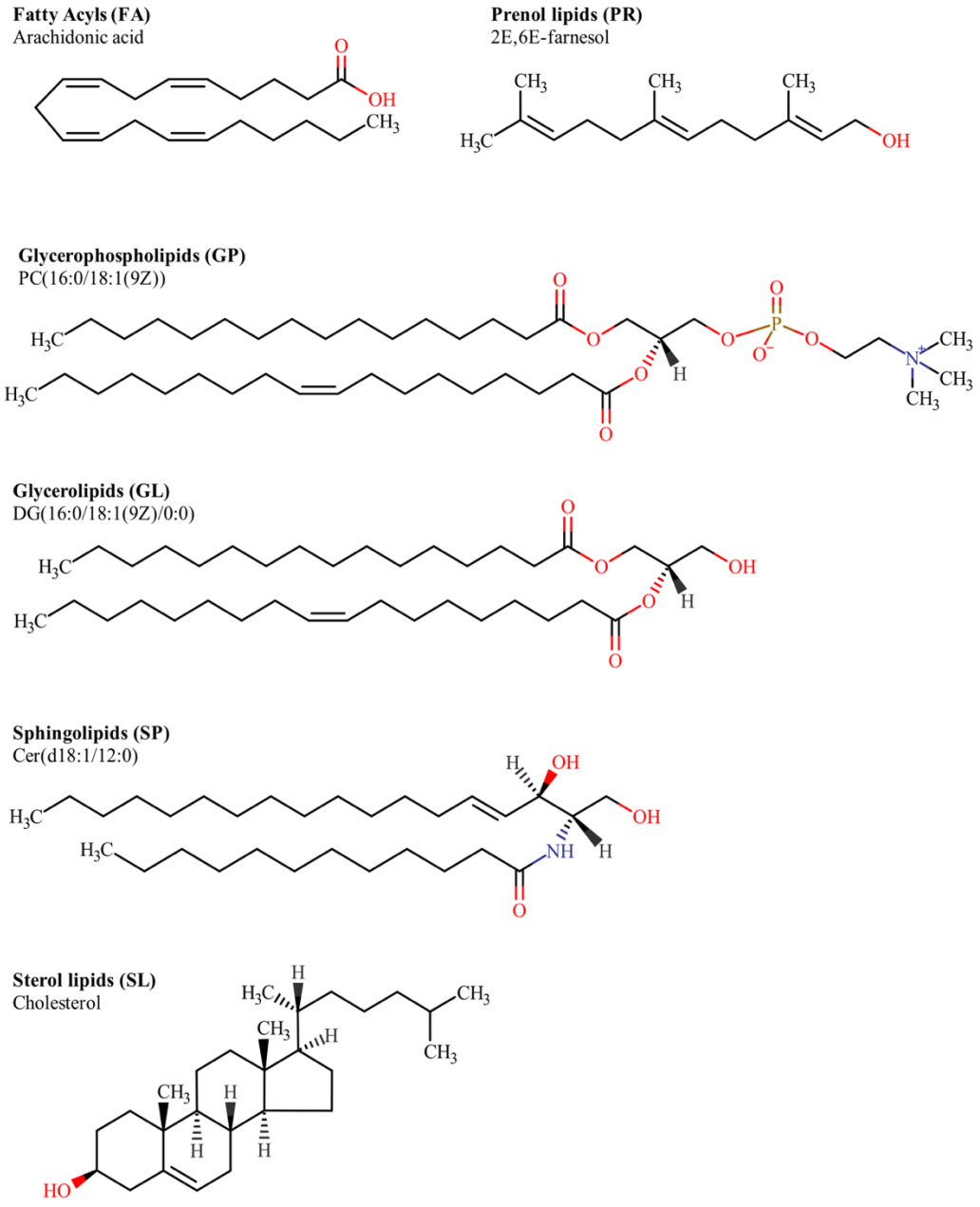


Figure 5.10 Representative lipid structures from the LIPID MAPS structure database for different lipid categories.

Lipid category and name are shown under each structure.

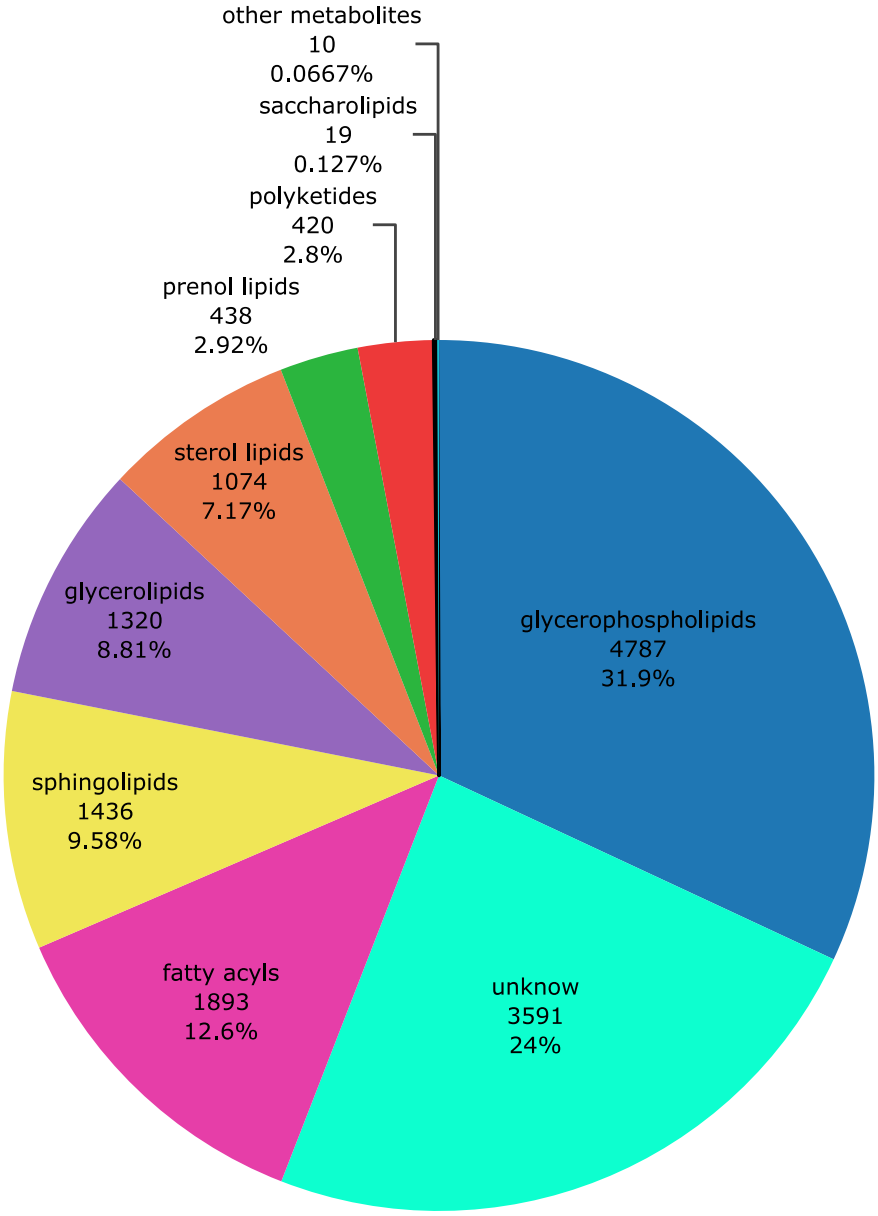


Figure 5.11 Global view of the lipidome of pRes Mϕ.

The pie chart shows the contribution of the different lipid categories to the total lipidome (number of lipid species putatively identified), values below each category represents the number of lipids putatively identified for each category, according to the classification on figure 6.4.

d) MDA with putative identifications reveals that lipid regulation by GATA6 varies depending on lipid category

Following the putative identification of the entire dataset, I mapped the lipid categories onto the results obtained in section 5.2.1b)i), in an effort to discern which lipid categories were responsible for the lipidomic variance between *Gata6-WT* and -KO^{mye} pResMϕ. The loadings in Figure 5.12 indicate that the observed separation between the two groups is mostly due to an increase in lipid levels, from different categories, in *Gata6-WT and* -KO^{mye} pResMϕ, evidenced by the higher numbers of features on the 2nd and 4th quadrant of the plot (positive PC1 values). Overall the results from the univariate and multivariate analysis seem to match, with the lipids identified as significantly altered being positioned on opposite ends on PC1.

Of interest 8.4 % of total identified lipids showed significant differences, however, in 2 of the 8 lipid categories (SL and GPL) the percentage of significantly altered lipids, in *Gata6*-KO^{mye}, was considerably higher (9 – 15 %). The lipid class where percentage change was lowest was FA (3 %), Figure 5.13.

Next, I will analyse the lipidomic differences by lipid category, which will allow me to investigate if which lipid species and pathways have been significantly altered by GATA6 deficiency.

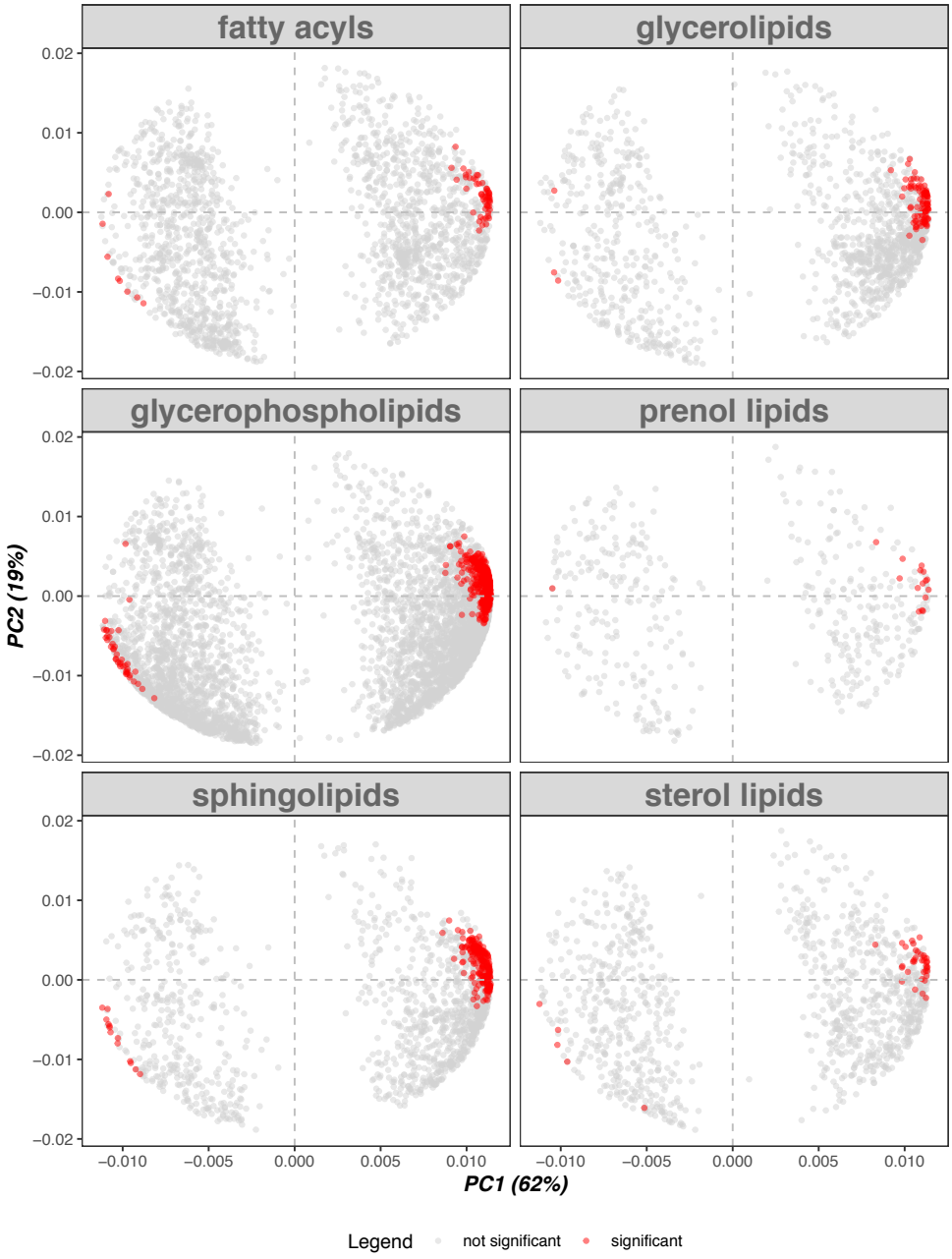


Figure 5.12 Mapping of multivariate data analysis with putative identifications reveals that lipid regulation by GATA6 varies depending on lipid category.

Figure shows loadings plots for PCA models by lipid category. Significantly altered lipids in *Gata6*-KO^{mye} pResMϕ are represented in red. Statistical significance was determined using the student t-test on log transformed data, with alpha=5.00 %. Each row was analysed individually, without assuming a consistent SD, and P-values were adjusted using B.H correction. Differential expression was classified as significant when P< 0.1 and |FC| > 1.5 (depicted in red). Note both fold changes and p-values are log transformed. The further its position away from the (0,0), the more significant the feature is.

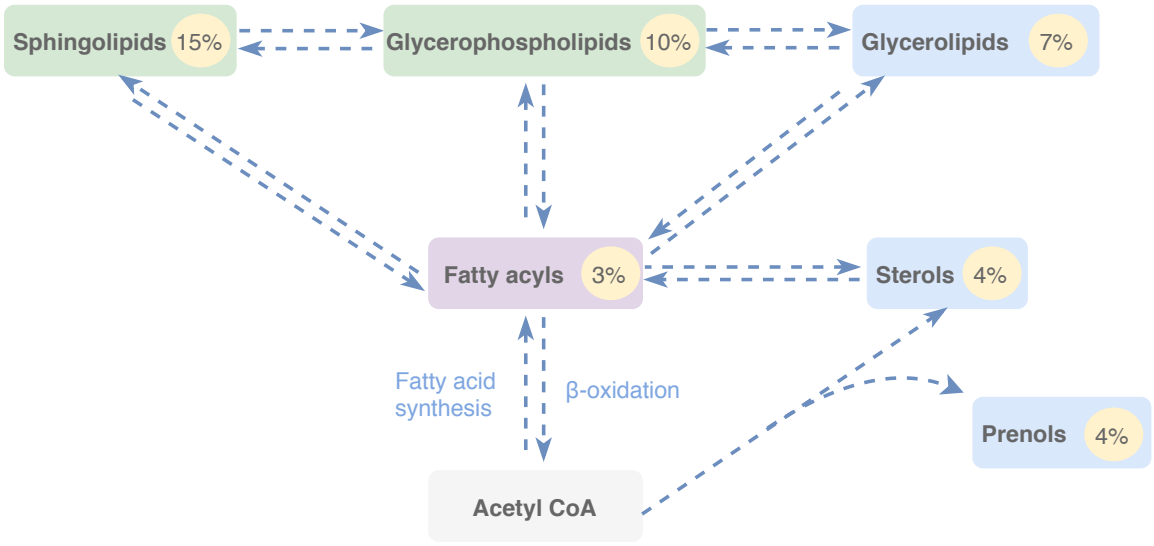


Figure 5.13 Interrelationships between cellular metabolic pathways.

Blue arrows denote multistep transformations among the major lipid categories and acetyl CoA. Values in yellow circles represent the percentage of analytes, putatively identified as belonging to each lipid category that were significantly changed in *Gata6*-KO^{mye} cells. Green rectangles represent the lipid categories that changes the most.

5.2.2 Untargeted analysis of lipid by category

The pResMϕ lipidome is comprised of hundreds of thousands of unique molecular species of lipids and metabolism in a given mammalian cell involves many thousands of these. In my attempt to define the differences between *Gata6-WT and* -KO^{mye} pResMϕ, I have focused on the putatively identified lipid categories that changed the most between the two cell types. A global overview of the changes in the murine Mϕ lipidome upon *Gata6*-KO^{mye} is depicted in Figure 5.13 which illustrates the interrelationships between major lipid categories, the percentage of analytes and those putatively identified as belonging to each lipid category, that were significantly changed in *Gata6*-KO^{mye} vs -WT cells. Based on the analysis peformed in the previous section, the 2 lipid categories with the highest percentages of significantly altered lipids, were selected for an in-depth look into which species are being altered as a consequence of *Gata6* KO (SL (15 %) and GPL (10%); represented in green, Figure 5.13).

Acetyl CoA is converted into fatty acids, sterols, and other intermediary metabolites in the six major lipid categories. Fatty acids (a subset of FA) are converted into bioactive eicosanoids (also a subset of FA) and more complex lipids such as GL, GPL, and SL.

Additionally, due to the central role of fatty acids in lipid metabolism (Figure 5.13), I will also investigate what molecular species of this class were affected by the downregulation of *Gata6*.

a) FA

The analysis performed in the previous section revealed that only 4 % of FA were significantly changed in in *Gata6*-KO^{mye} vs control cells. FA (including fatty acids) are a lipid category comprising a diverse group of molecules synthesized by chain-elongation of an acetyl-CoA primer with malonyl-CoA or methylmalonyl-CoA groups, and therefore is one of the most fundamental categories of biological lipids (see Chapter 1, Table 1.2 for the complete list of lipid classes).

First, I selected the features identified by the volcano (Figure 5.14) plot as significant and used an in-house generated retention time and *m/z* database which identifies around 1,000 lipids known to be present in mammalian cells for putative identification (level three identification according to the metabolomics standard initiative[206]) (Figure 5.15). This analysis revealed that the main class with the higher percentage of significantly changed lipid levels was fatty acyl glycosides, simple glycolipids generally constituted of a glycosyl moiety (one or several units) linked to the hydroxyl group of a hydroxy fatty acid or to one carboxyl group of a normal fatty acid (ester linkage). Several more complex glycosides are also formed by amide linkage between a glycosyl moiety and a fatty acid.

Next, I considered free fatty acids known to be present in pResMϕ, due to their relevance for other lipid categories.

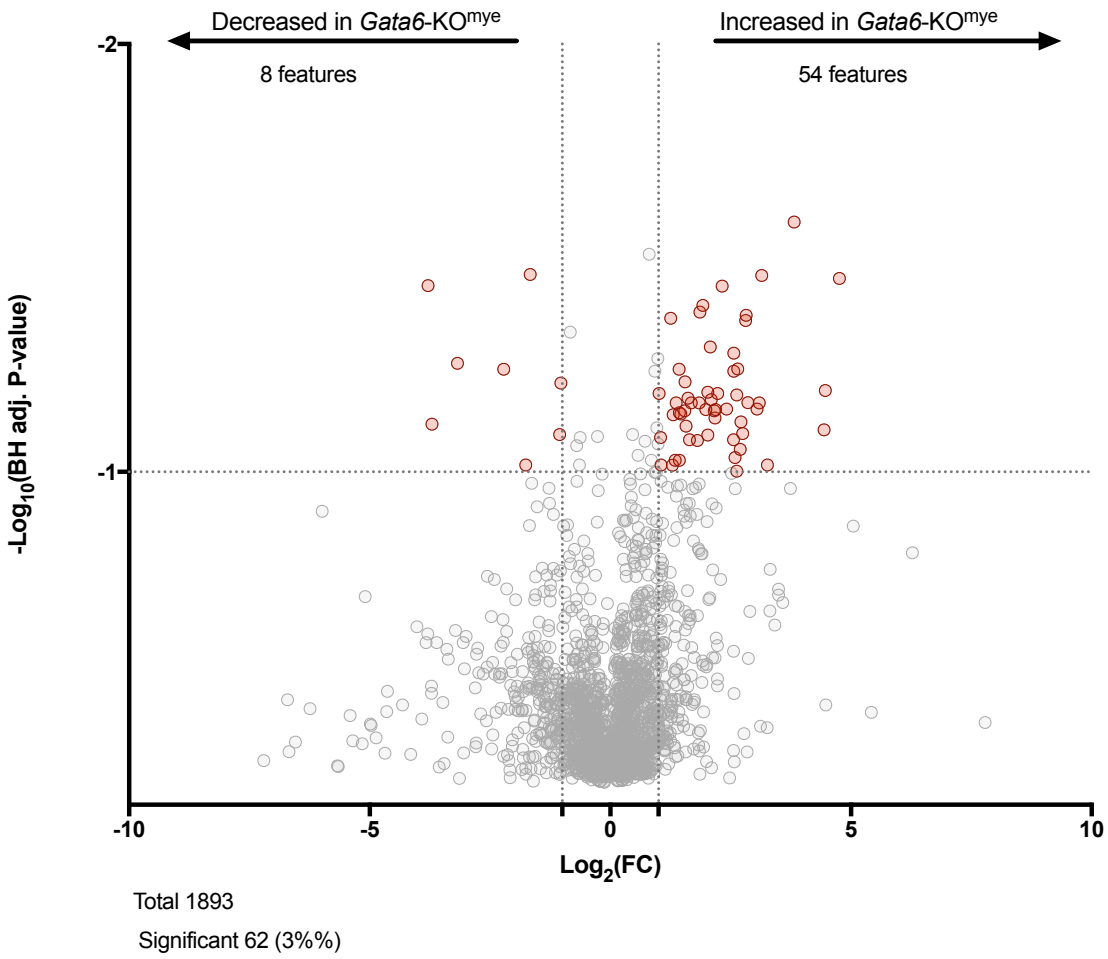


Figure 5.14 3% of FA vary significantly between *Gata6*-WT and -KO pResMϕ.

Mϕ lipid extracts from *Gata6*-WT and -KO pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. Volcano plots representing fold-change and p-value of lipids in *Gata6*-WT relative to -KO pResMϕ. Lipids were putatively identified using the MSSearch program, which used the LIPIDMAPS database to attribute a lipid category to each lipid. Significantly changed lipid levels are indicated by red (3 %) (of which 87 % were increased in GATA6 deficient cells), and unchanged levels by grey (97 %). Statistical significance was determined using the student t-test on log transformed data, with alpha=5.00%. Each row was analysed individually, without assuming a consistent SD, and P-values were adjusted using B.H correction. Differential expression was classified as significant when $P < 0.1$ and $|FC| > 1.5$ (depicted in red). Note both fold changes and p-values are log transformed. The further its position away from the (0,0), the more significant the feature is.

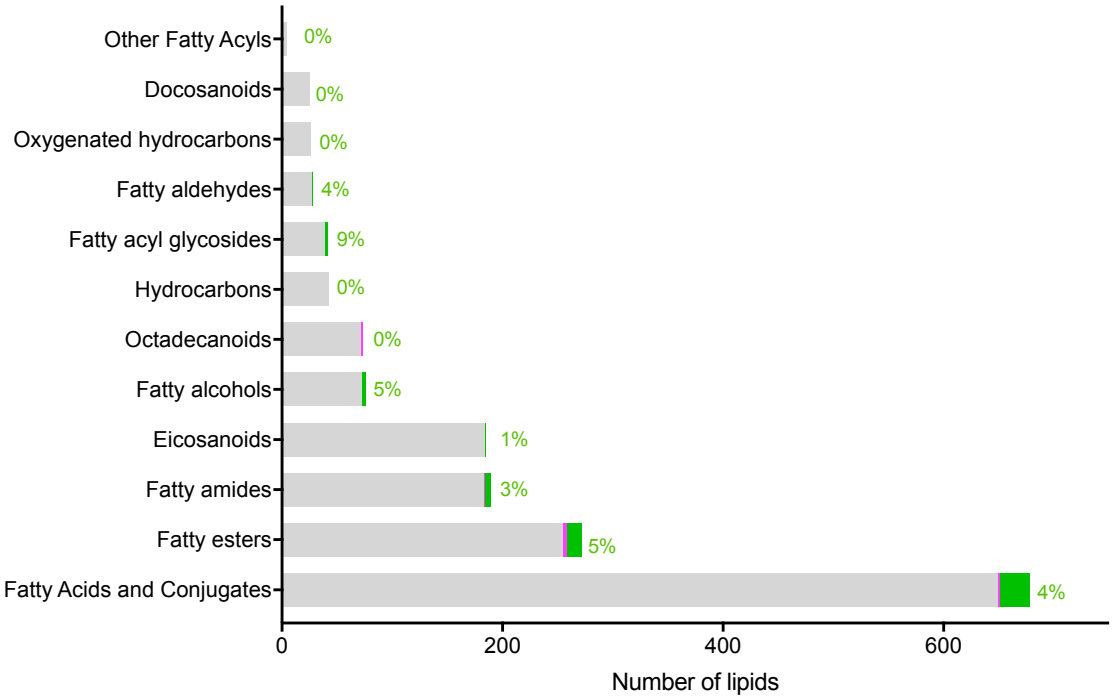


Figure 5.15 9% of fatty acyl glycosides increase significantly in *Gata6*-KO^{mye} pResMϕ.
Histogram shows the number of FA species by lipid class. Grey corresponds to the number of lipids that did not change significantly; magenta, lipids that decreased significantly; green, lipids that increased significantly. Significance was determined using the student t-test on log transformed data, corrected using B.H. with alpha=10.00 %. Differential expression was classified as significant when $P < 0.1$ and $|FC| > 1.5$.

i) Free fatty acids

Knowledge of the fatty acid composition of lipid classes is an important factor in the investigation of metabolism. Thus, I compiled a list of 32 putatively identified fatty acids, known to be present in Mϕ-like cells[95], and split them into four groups, based on the degree of saturation and position of the double bonds. These lipids were identified based on their accurate *m/z* and retention time. The four groups were saturated fatty acids and ω-3, -6, and -9 mono- and polyunsaturated fatty acids (Table 5.2). The designation ω-fatty acid refers to the position of the first double counted from the ω carbon (or n position), for example, the first double bond (counted from the ω carbon) in alpha linolenic acid (ALA) (18:3 (n-3)), represented in Figure 5.16 is positioned on the 3rd carbon, thus ALA is an ω-3 polyunsaturated fatty acid.

Fatty acids changed as a consequence of *Gata6*-KO. Levels of ω-6 and ω-9 unsaturated free fatty acids were increased in *Gata6*-KO^{mye} pResMϕ. Out of 31 molecular species, 4 were significantly increased in *Gata*6-KO^{mye}. Additionally, there was a trend for the fatty acids with a higher FC corresponded to lipids with a higher degree of unsaturation, which can be seen more clearly in the pathways depicted in Figure 5.18 and Figure 5.19.

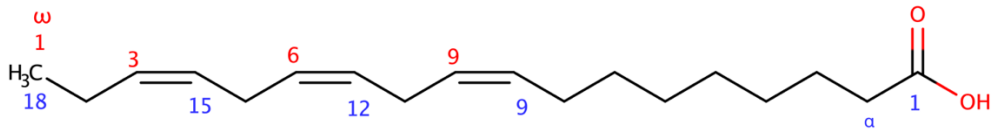


Figure 5.16 Chemical structure of alpha-linolenic acid (ALA), an essential omega−3 fatty acid.

18:3Δ9c,12c,15c, which means a chain of 18 carbons with 3 double bonds on carbons numbered 9, 12, and 15.

Table 5.2 List of putative fatty acids found in pResMϕ.

Saturated fatty acids

Common name	Lipid name	Chemical name
Lauric acid	C12:0	dodecanoic acid
Myristic acid	C14:0	tetradecanoic acid
Pentadecylic acid	C15:0	pentadecanoic acid
Palmitic acid	C16:0	hexadecanoic acid
cis-9-palmitoleic acid	C16:1n-7	9Z-hexadecenoic acid
Margaric acid	C17:0	heptadecanoic acid
10Z-heptadecenoic acid	C17:1n-7	10Z-heptadecenoic acid
Stearic acid	C18:0	octadecanoic acid
Arachidic acid	C20:0	eicosanoic acid
Behenic acid	C22:0	docosanoic acid
Tricosanoic acid	C23:0	tricosanoic acid
Lignoceric acid	C24:0	tetracosanoic acid
Cerotic acid	C26:0	hexacosanoic acid

Omega-3 fatty acids, polyunsaturated

Common name	Lipid name	Chemical name
Hexadecatrienoic acid (HTA)	16:3 (n-3)	*all-cis* 7,10,13-hexadecatrienoic acid
Alpha-linolenic acid (ALA)	18:3 (n-3)	*all-cis* -9,12,15-octadecatrienoic acid
Stearidonic acid (SDA)	18:4 (n-3)	*all-cis* -6,9,12,15,-octadecatetraenoic acid
Eicosatrienoic acid (ETE)	20:3 (n-3)	*all-cis* -11,14,17-eicosatrienoic acid
Eicosapentaenoic acid (EPA, Timnodonic acid)	20:5 (n-3)	*all-cis* -5,8,11,14,17-eicosapentaenoic acid
DTrE	22:3 (n-3)	*all-cis* -3,6,9-docosatrienoic acid
Docosapentaenoic acid (DPA, Clupanodonic aci	22:5 (n-3)	*all-cis* -7,10,13,16,19-docosapentaenoic acid
Docosahexaenoic acid (DHA, Cervonic acid)	22:6 (n-3)	*all-cis* -4,7,10,13,16,19-docosahexaenoic acid

Omega-6 fatty acids, polyunsaturated

Common name	Lipid name	Chemical name
Linoleic acid	18:2 (n-6)	*all-cis* -9,12-octadecadienoic acid
Gamma-linolenic acid (GLA)	18:3 (n-6)	*all-cis* -6,9,12-octadecatrienoic acid
Eicosadienoic acid	20:2 (n-6)	*all-cis* -11,14-eicosadienoic acid
Dihomo-gamma-linolenic acid (DGLA)	20:3 (n-6)	*all-cis* -8,11,14-eicosatrienoic acid
Arachidonic acid (AA)	20:4 (n-6)	*all-cis* -5,8,11,14-eicosatetraenoic acid
Docosadienoic acid	22:2 (n-6)	*all-cis* -13,16-docosadienoic acid
Adrenic acid	22:4 (n-6)	*all-cis* -7,10,13,16-docosatetraenoic acid
Docosapentaenoic acid (Osbond acid)	22:5 (n-6)	*all-cis* -4,7,10,13,16-docosapentaenoic acid

Omega-9 fatty acids, mono- and polyunsaturated

Common name	Lipid name	Chemical name
Oleic acid	18:1 (n-9)	*cis* -9-octadecenoic acid
Mead acid	20:3 (n-9)	*all-cis* -5,8,11-eicosatrienoic acid
Erucic acid	22:1 (n-9)	*cis* -13-docosenoic acid
Nervonic acid	24:1 (n-9)	*cis* -15-tetracosenoic acid

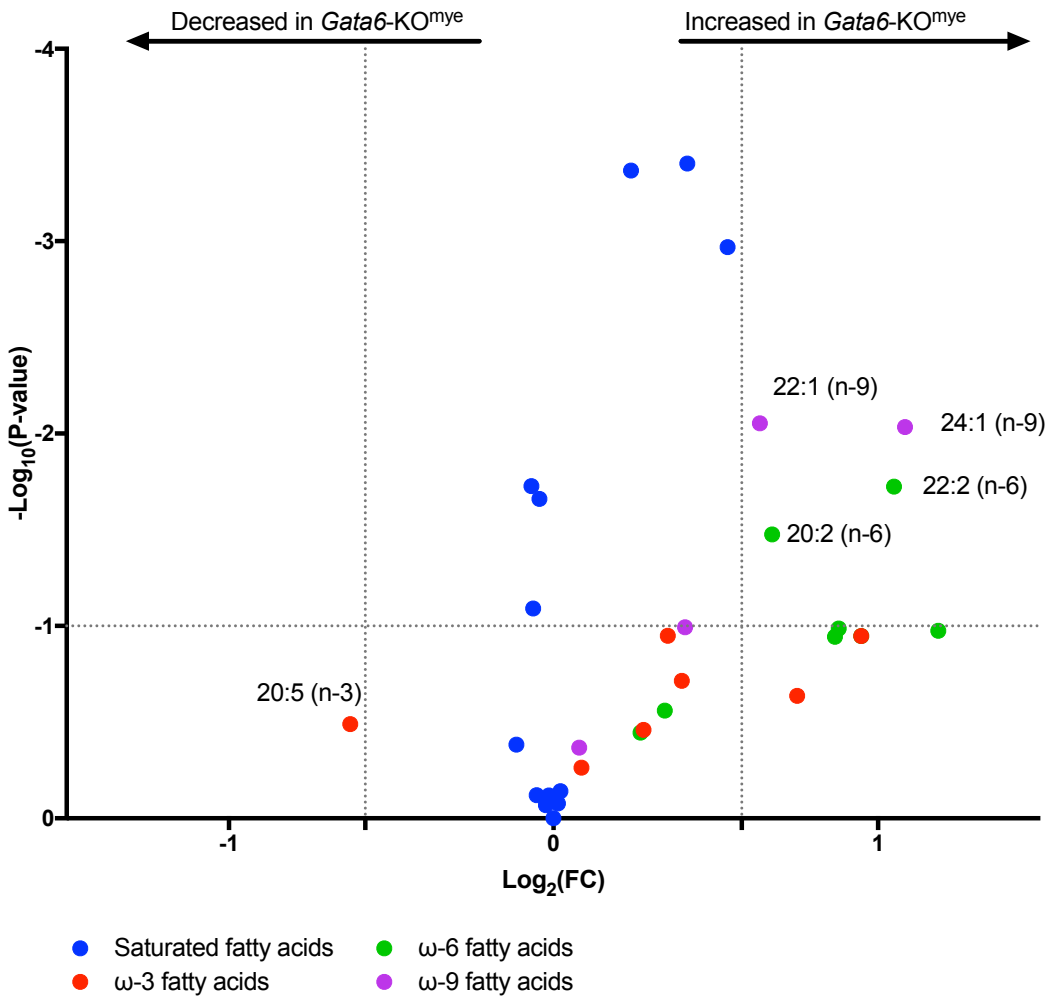


Figure 5.17 ω-6 and -9 FA increase significantly between *Gata6*-WT and -KO pResMϕ.
Mϕ lipid extracts from *Gata6*-WT and -KO pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. Volcano plots representing fold-change and p-value of free FA in *Gata6*-WT relative to -KO pResMϕ. Lipids were putatively identified using an inhouse database, which used the *m/z* and RT to attribute a putative identification to each lipid. Statistical significance was determined using the student t-test, with alpha=10.0 %. Each row was analysed individually, without assuming a consistent SD. Differential expression was classified as significant when $P < 0.05$ and $|FC| > 1.5$ (depicted in red). Note both fold changes and p-values are log transformed. The further its position away from the (0,0), the more significant the feature is.

Pathway analysis

The pathway analysis in in Figure 5.18 and Figure 5.19 illustrates the changes in metabolites and transcripts leading from alpha-linolenic acid and linoleic acid to the ω-3 and ω-6 unsaturated fatty acids (Figure 5.18), and from lauric acid to the saturated and ω-9 unsaturated fatty acids in response to *Gata6*-KO^{mye} in pResMϕ.

Multiple genes encoding fatty acid elongases (Elovl) and desaturases (Fads) were down-regulated in to *Gata6*-KO^{mye}. However, these decreases did not correlate with a higher abundance of free fatty acid levels. The levels of these lipids can also be related to the activity of acyl transferases, which remodel them into complex lipids (these will be discussed in the next sections).

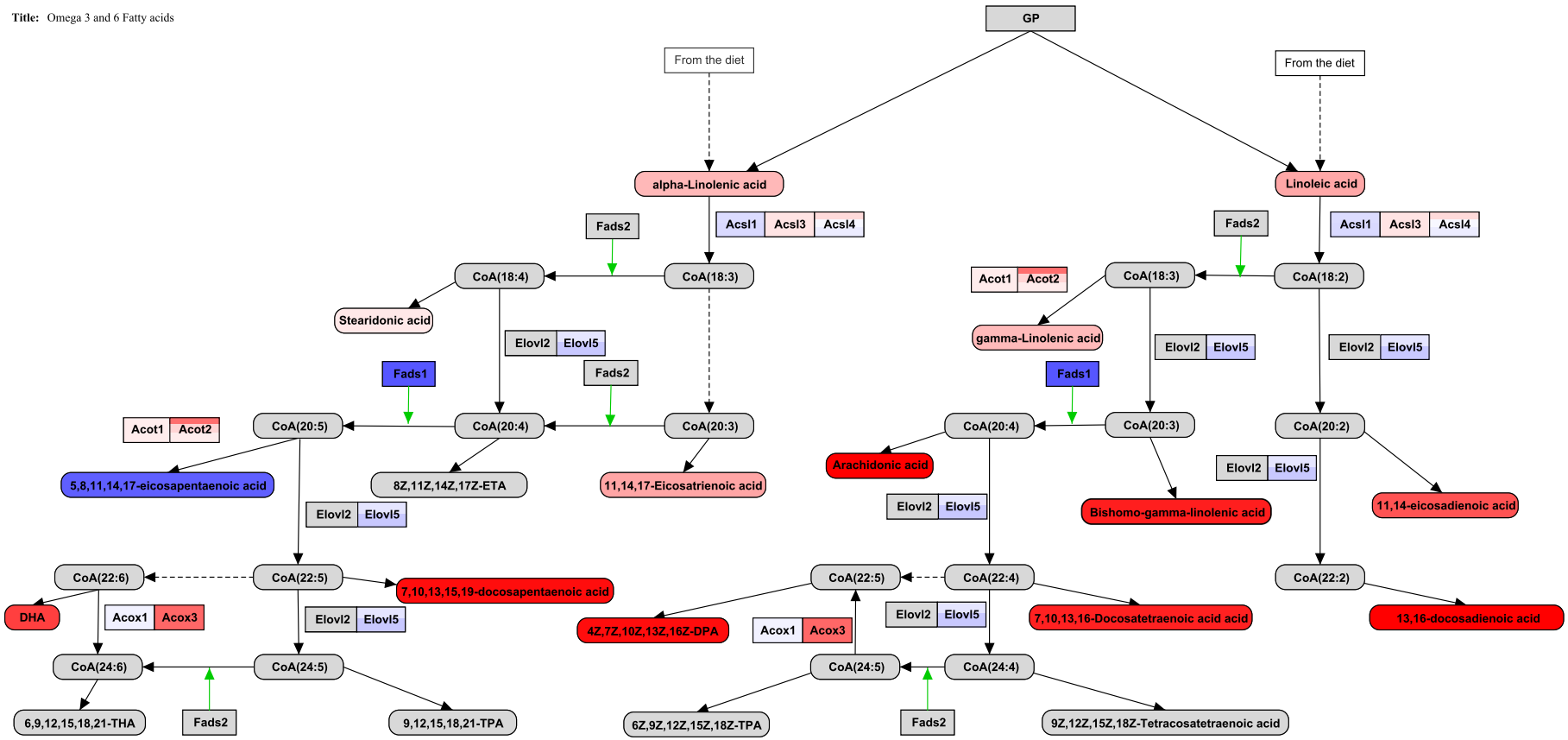


Figure 5.18 Changes in metabolites and transcripts leading from alpha-linolenic acid and linoleic acid to the ω-3 and ω-6 unsaturated fatty acids in response to *Gata6*-KO^{mye} in pResMϕ.

Rectangular boxes represent genes coding for enzymes that catalyse the indicated conversion, while boxes with rounded edged represent lipid metabolites. The colour in each lipid and gene represent the Log_2(FC). Lipid and mRNA levels were obtained as described in Chapter 2.

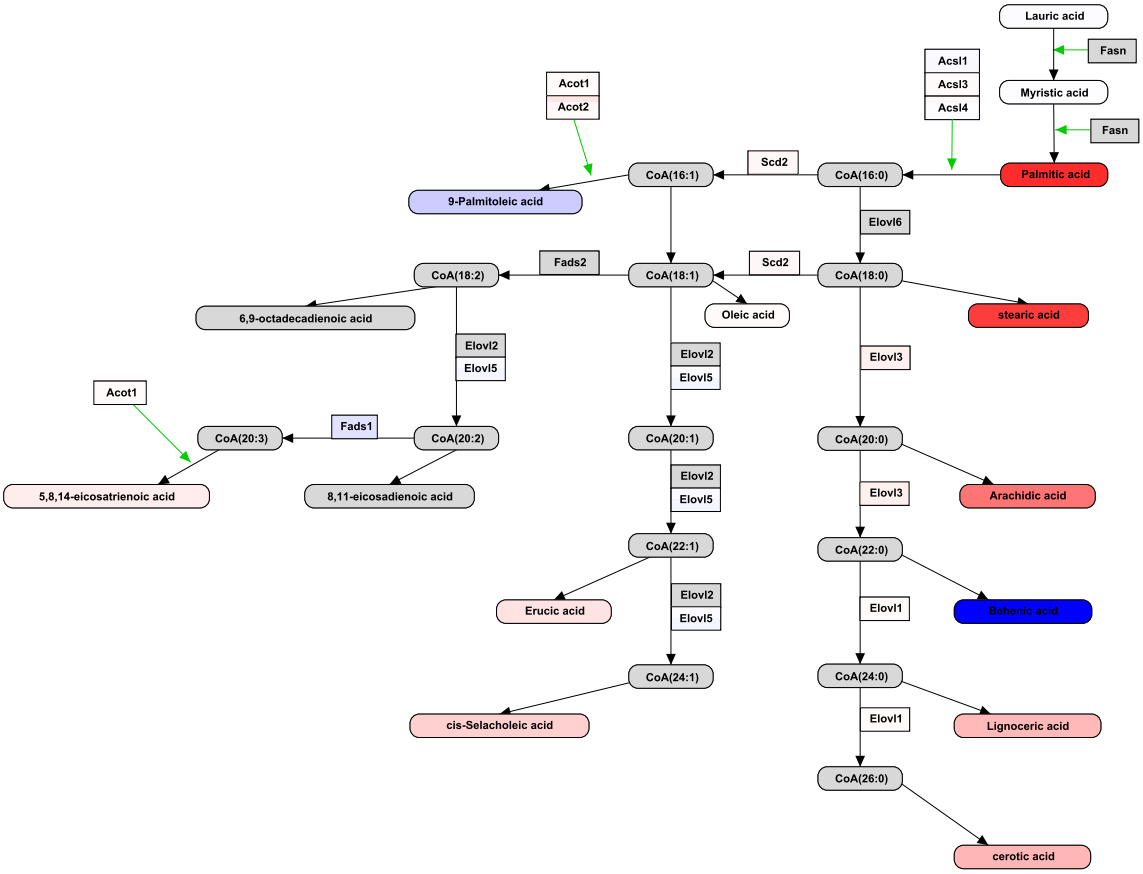


Figure 5.19 Changes in metabolites and transcripts leading from lauric acid to the saturated and ω-9 unsaturated fatty acids in response to *Gata6*-KO^{mye} in pResMϕ.

Rectangular boxes represent genes coding for enzymes that catalyse the indicated conversion, while boxes with rounded edged represent lipid metabolites. The colour in each lipid and gene represent the Log_2(FC). Lipid and mRNA levels were obtained as described in Chapter 2.

b) SL

Putatively identified SL (Figure 5.20), presented the highest percentage of significantly altered lipids (15 %), of which 94% were increased (Figure 5.21). Complex SL are comprised of a sphingoid long-chain base (LCB) linked via the amide bond to an N-acylated fatty acid to yield Cer (the basic unit of all SL) and a polar head group (Figure 5.20)[128,130,207,208]. As shown in Figure 5.20, the high structural multiplicity of SL is attributed to (a) the length of the sphingoid LCB (the most common chain being 18 carbon atoms) and the N-acylated fatty acid chain (often ranging from 14 to 36 carbon atoms)[128,130], (b) the degree of hydroxylation, and number and position of double bond in the LCB, (c) the saturation and hydroxylation status of the fatty acid, and (d) the substituent as the head group at position 1 of the basic Cer (which contains a simple hydroxyl group) [128,130]. Consequently, the resulting sphingolipidome is extremely complex, encompassing up to thousands of sphingoid bases and their derivatives. It is the structural diversity of SL that is believed to enable their diverse cellular functions including those related to immunity. There is a convenient short-hand nomenclature to describe SL with structural features[128]. For example, d18:1/C16:0 refers to a Cer consisting of So, which has an 18 carbon chain with 1 double bond (d18:1), and an amide-linked C16 fatty acid chain with 0 double bond (C16:0) (Figure 5.20).

Here, I selected the significantly changed features putatively identified as SL and attributed classes to each of the lipid features, as described in the previous section. This analysis revealed that the classes with the highest percentage of matches were also the ones that changed the most upon *Gata6* KO, neutral SL, Cer and phosphosphingolipids (Figure 5.22). These classes showed increased lipid levels in 9 – 28 % of all putatively identified features (phosphosphingolipids 9%, Cer 13 %, and neutral SL 28 %) and I will look into them in more detail in the next section.

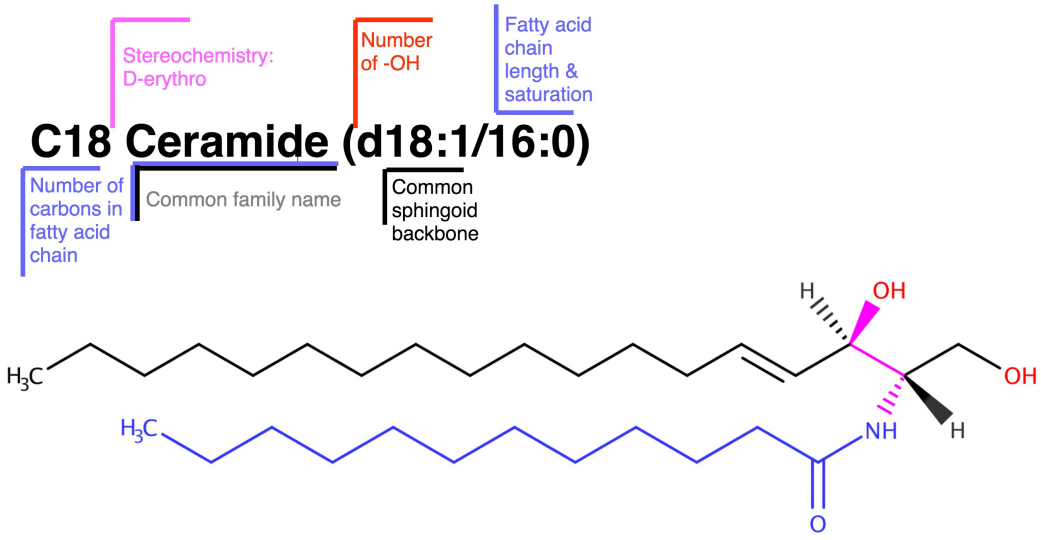


Figure 5.20 SL structure and nomenclature.

All the structural variables are highlighted. Ceramide (Cer) is the fundamental unit of all complex SL. The Cer consists of two structural moieties: the sphingoid long-chain base (LCB) depicted in black, and the fatty acid chain linked via an amide bond, blue. The typical LCB has a chain length of 18 carbons, which may be hydroxylated at 4-position, or have a double bond at the 4 or 8 carbon. The fatty acid chain may be hydroxylated at the α-position, and/or have a double bond at ω9-position. The fatty acid chain length may vary from 14 to 36 (if >20, it is referred to as very long-chain fatty acid, *i.e.*, VLCFA). The structurally diverse Cer species can be converted to more complex SL via substitution of the hydroxyl group at the 1-position of the LCB. Additional sugar residues may be further added, resulting in more complex SL.

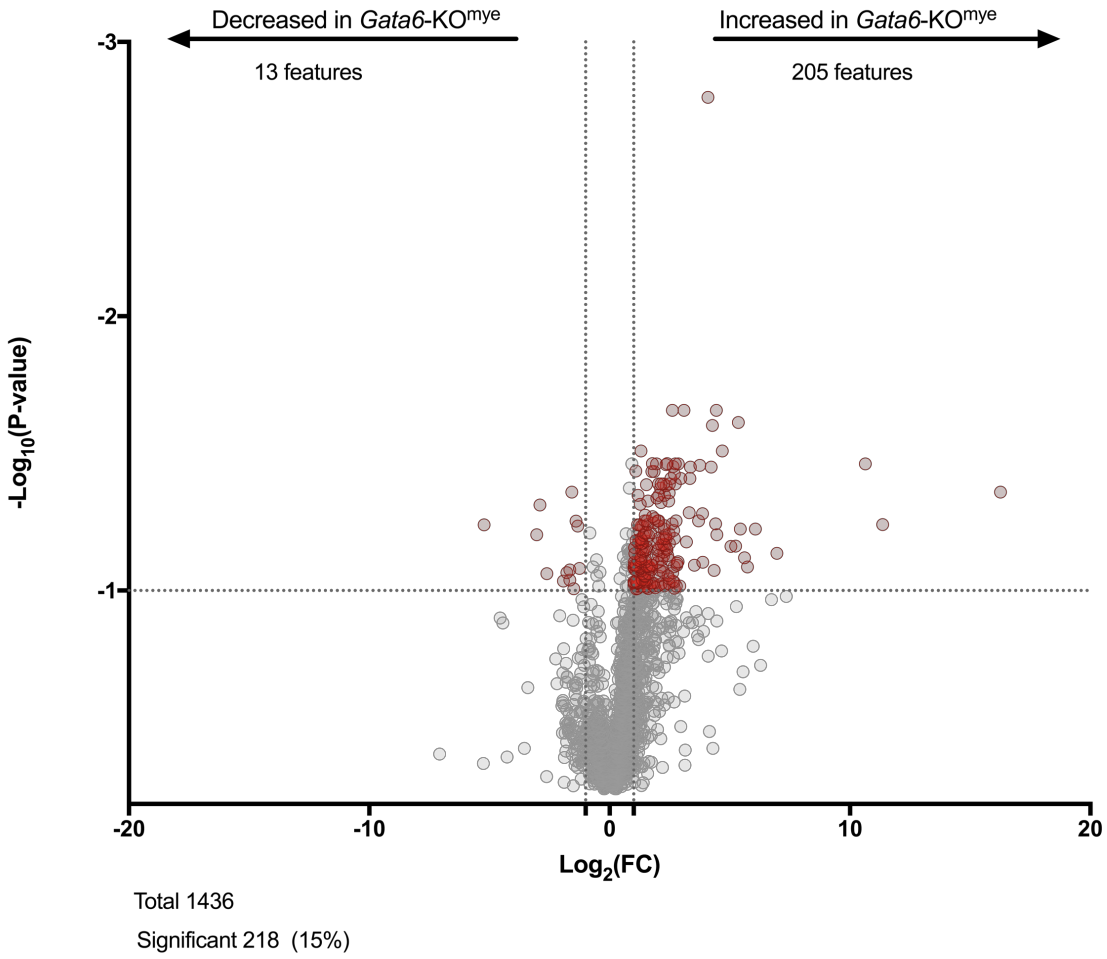


Figure 5.21 15 % of SL vary significantly between *Gata6*-WT and -KO pResMϕ.

Mϕ lipid extracts from *Gata6*-WT and -KO pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. Volcano plots representing fold-change and p-value of lipids in *Gata6*-WT relative to -KO pResMϕ. Lipids were putatively identified using the MSSearch program, which used the LIPIDMAPS database to attribute a lipid category to each lipid. Significantly changed lipid levels are indicated by red (15 %) (of which 94 % were increased in GATA6 deficient cells), and unchanged levels by grey (85 %). Statistical significance was determined using the student t-test on log transformed data, with alpha=10.00 %. Each row was analysed individually, without assuming a consistent SD, and P-values were adjusted using B.H correction. Differential expression was classified as significant when $P < 0.1$ and $|FC| > 1.5$ (depicted in red). Note both fold changes and p-values are log transformed. The further its position away from the (0,0), the more significant the feature is.

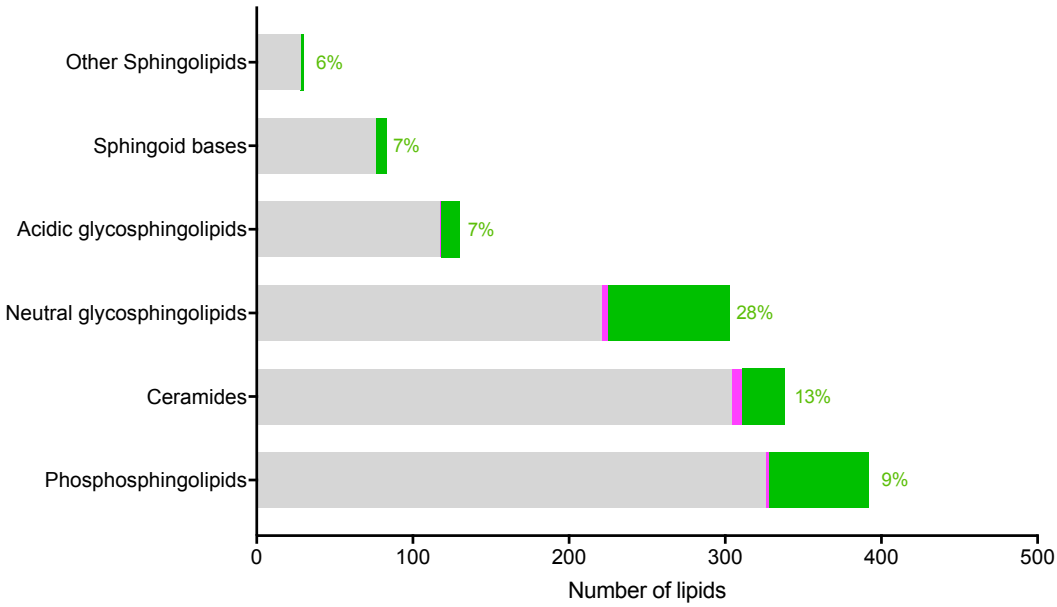


Figure 5.22 28 % of neutral SL increase significantly in *Gata6*-KO^{mye} pResMϕ.

Histogram shows the number of SL species by lipid class. Grey corresponds to the number of lipids that did not change significantly; magenta, lipids that decreased significantly; green, lipids that increased significantly. Significance was determined using the student t-test on log transformed data, corrected using B.H. with alpha=10.00 %. Differential expression was classified as significant when $P < 0.1$ and $-|FC| > 1.5$.

i) Neutral SL, Cer and phosphosphingolipids

I considered a group of 50 putatively identified neutral SL, Cer and phosphosphingolipids known to be present in pResMϕ, in order to establish common structural features amongst the significantly increased lipids. I found that out of the 50, 30 were significantly increased in the *Gata6*-KO pResMϕ (Figure 5.23). Of these, the highest fold changes corresponded to long chain HexCer and SM (Figure 5.24).

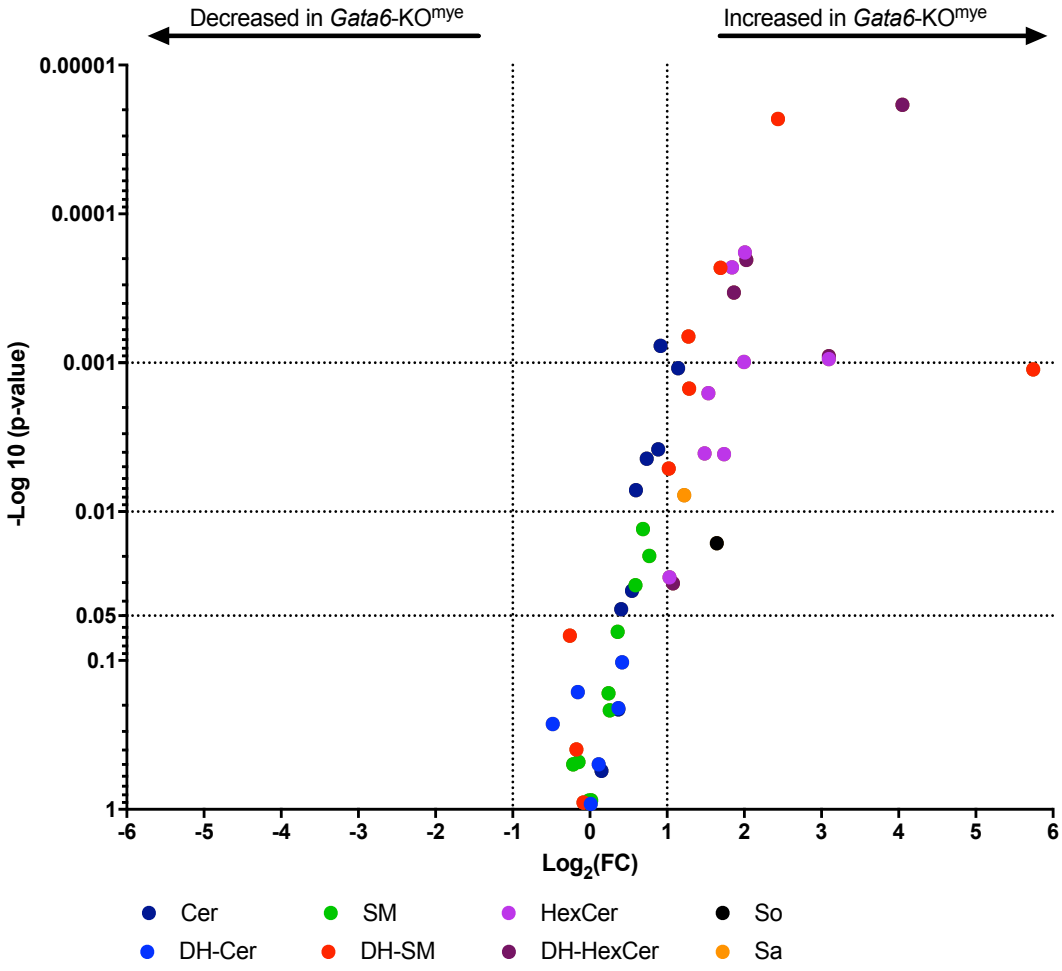


Figure 5.23 SM and HexCer increase significantly in *Gata6*-KO^{mye} pResMϕ.

Mϕ lipid extracts from *Gata6*-WT and -KO pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. Volcano plot represents the fold-change and p-value of SL in *Gata6*-WT relative to -KO pResMϕ. Lipids were putatively identified using an inhouse database, which used the *m/z* and RT to attribute a putative identification to each lipid. Statistical significance was determined using the student t-test, with alpha=10.00 %. Each row was analysed individually, without assuming a consistent SD. Differential expression was classified as significant when $P < 0.05$ and $|FC| > 1.5$ (depicted in red). Note both fold changes and p-values are log transformed. The further its position away from the (0,0), the more significant the feature is.

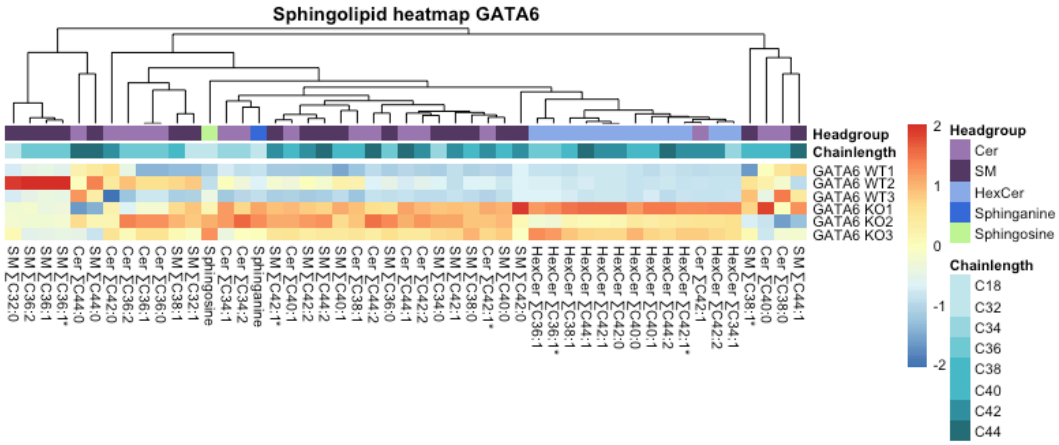


Figure 5.24 Effect of *Gata6*-KO on SL composition in pResMϕ.

Heatmaps were generated with intensity values using the pheatmap package in R. Hierarchical cluster analysis for SL present in *Gata6*-KO^{mye} and –WT pResMϕ. Intensity levels are represented by a colour gradient ranging from blue (decrease in response) to white (no change) to red (increase in response). Each lipid was annotated, using a colour code, according to the number of double bonds and chain length; the variations in signal intensity are represented in colour range, from blue to red (from the least to the most intense signals), and clustered by similarity in overall response to *Gata6*-KO.

c) GPL

GPL was the lipid category with the most lipid species putatively identified (37% of all lipid features). With 10 % of these molecules significantly altered by *Gata6* KO. GPL, are ubiquitous in nature and are key structural components of cell membranes. They are precursors of intracellular signalling molecules that modulate membrane trafficking and signal transduction. The GPL lipid category can be subdivided into several distinct classes, based on the nature of the polar headgroups (Figure 5.25).

I selected the features identified as GPL and subdivided them into lipid classes (Figure 5.26). This analysis revealed that the class with the highest percentage of significantly changed lipid levels was glycerophosphoglycerols[6], with 16% of lipids significantly increased.

Next, I selected a group of 26 PG with increasing chain lengths and degrees of saturation, in order to establish common structural features amongst the significantly increased lipids. I found that out of the 26, 19 were significantly increased in the *Gata6*-KO pResMϕ (Figure 5.28) with a clear trend for the levels of shorter chain PG to be more increased than longer chain PG.

[6] There were two other lipid classes with higher percentages of significantly increased lipids, however there were less than 10 lipids in each class and thus they were not considered for further analysis.

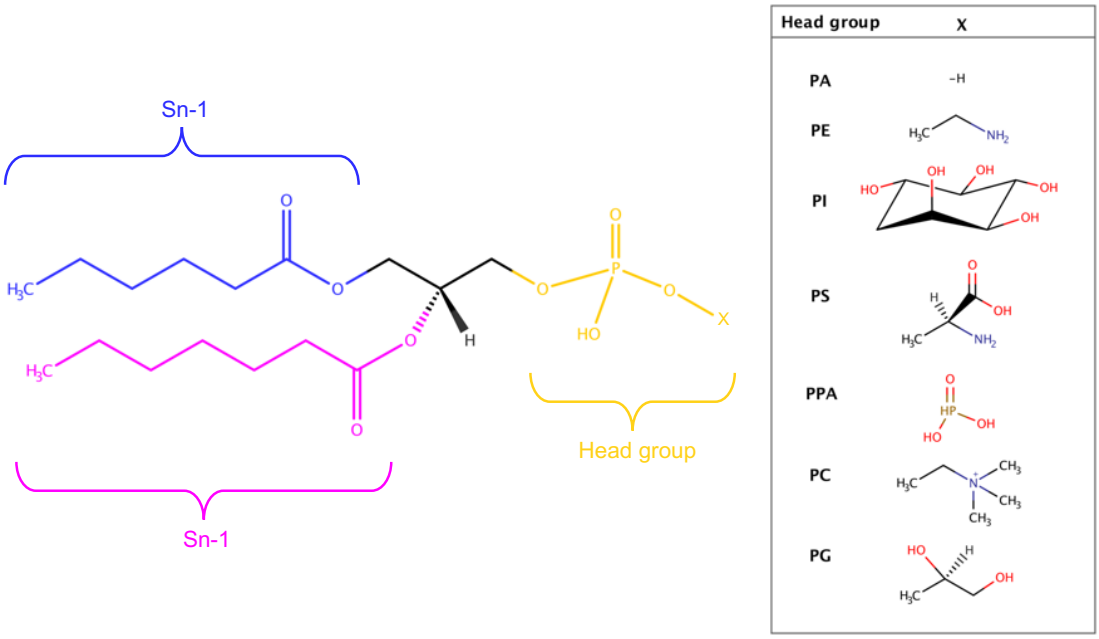


Figure 5.25 Structures of major GPL.

Head groups legend: PA = Phosphate, PPA = Pyrophosphate, PE = Phosphoethanolamine, PC = Phosphocholine, PS = Phosphoserine, PG = Phosphoglycerol, and PI = Phosphoinositol.

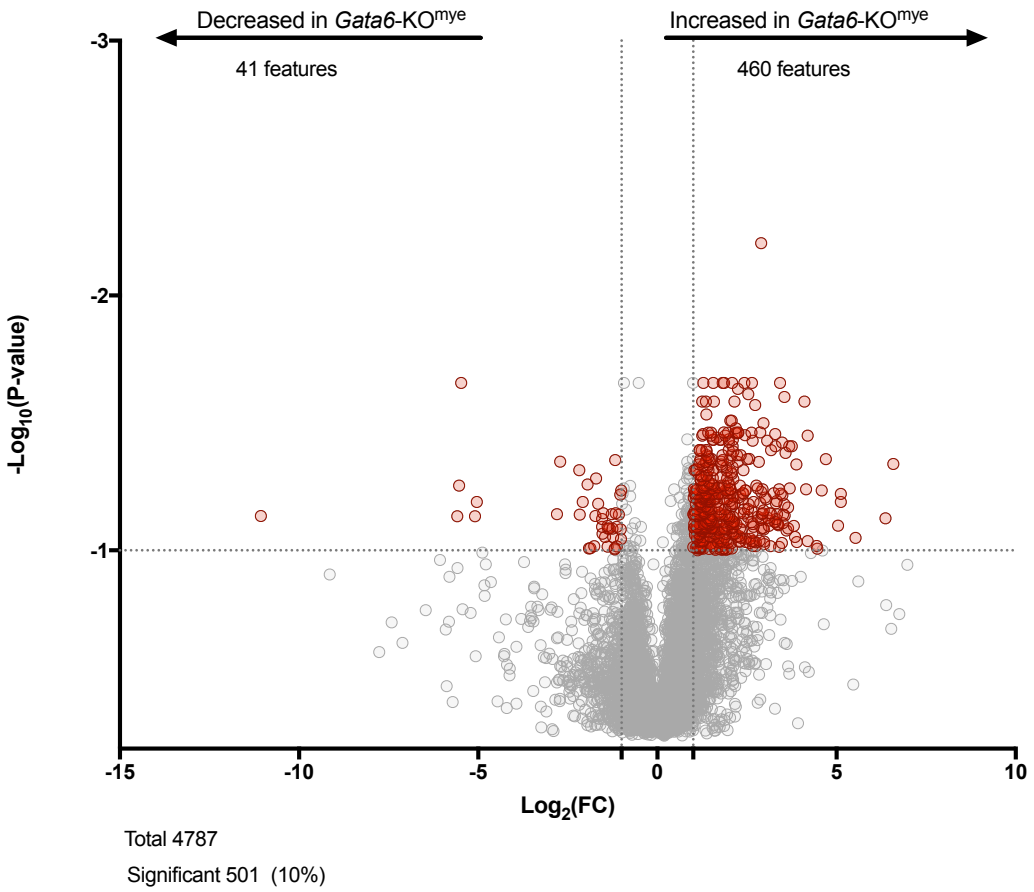


Figure 5.26 10 % of GPL vary significantly between *Gata6*-WT and -KO pResMϕ.

Mϕ lipid extracts from *Gata6*-WT and -KO pResMϕ were analysed by LC-FTMS on the Orbitrap Elite, at 60,000 resolution, then processed using XCMS. Volcano plots representing fold-change and p-value of GPL in *Gata6*-WT relative to -KO pResMϕ. Lipids were putatively identified using the MSSearch program, which used the LIPIDMAPS database to attribute a lipid category to each lipid. Significantly changed lipid levels are indicated by red (10 %) (of which 92 % were increased in GATA6 deficient cells), and unchanged levels by grey (90 %). Statistical significance was determined using the student t-test on log transformed data, with alpha=5.00 %. Each row was analysed individually, without assuming a consistent SD, and P-values were adjusted using B.H correction. Differential expression was classified as significant when $P < 0.1$ and $|FC| > 1.5$ (depicted in red). Note both fold changes and p-values are log transformed. The further its position away from the (0,0), the more significant the feature is.

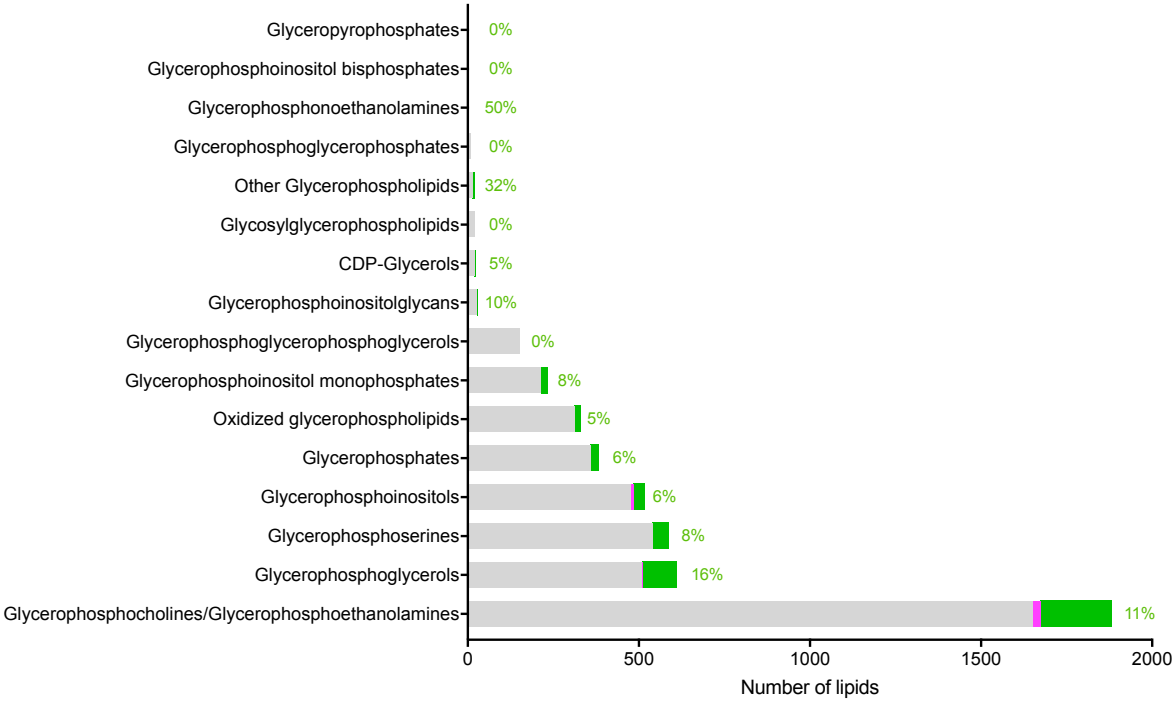


Figure 5.27 16 % of PG increase significantly in *Gata6*-KO^{mye} pResMϕ.

Histogram shows the number of GPL species by lipid class. Grey corresponds to the number of lipids that did not change significantly; magenta, lipids that decreased significantly; green, lipids that increased significantly. Significance was determined using the student t-test on log transformed data, corrected using B.H. with alpha=10.00 %. Differential expression was classified as significant when $P < 0.1$ and $|FC| > 1.5$.

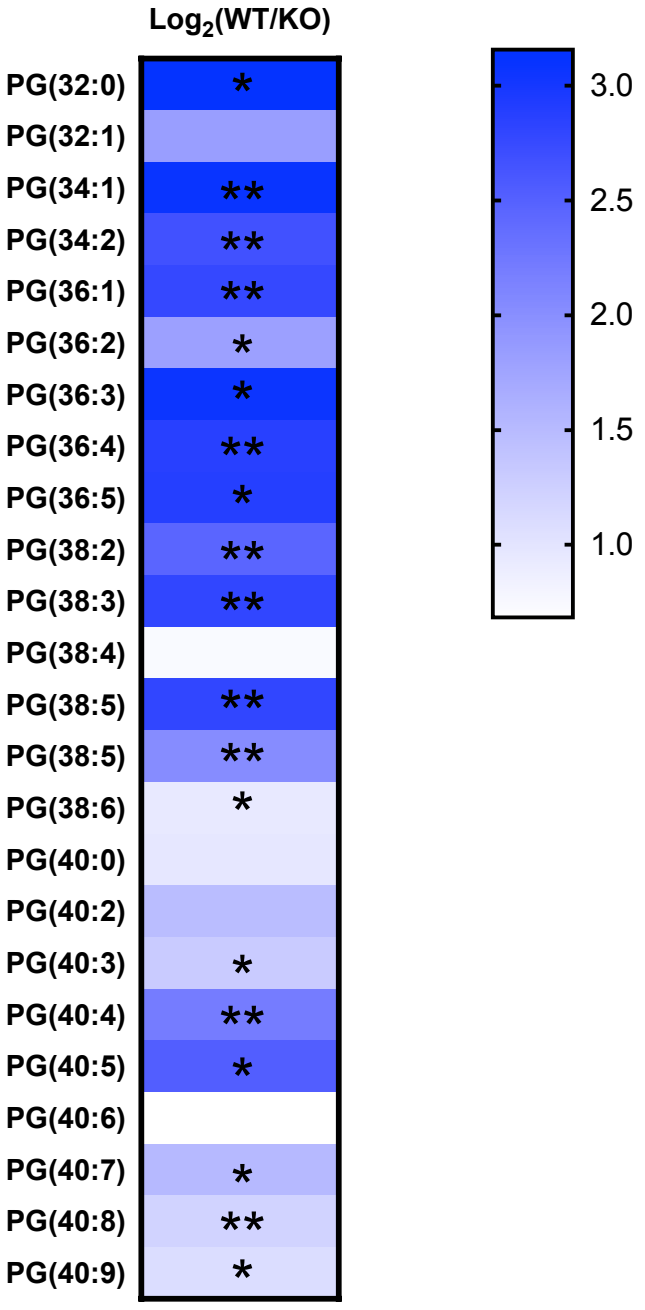


Figure 5.28 Alterations in the levels of PG during GATA KO.

The colour code mirrors the log_2(FC), the colour gradient from white to blue represents an increase in the lipid levels. The significances of the differences (*p < 0.05, **p < 0.005, ***p < 0.0005; t-test) are marked.

5.3 Discussion

In this Chapter, an LC-MS based lipidomics approach was used in an effort to elucidate if the transcription factor GATA6 regulates lipid metabolism in murine pResMϕ.

My lipid analysis strategy greatly facilitated interpretation of *Gata6*-WT and -KO^{mye} pResMϕ lipidomic profiles allowing me to identify several trends in the regulation of lipid levels. The power analysis performed in the beginning of this Chapter revealed that my experimental design is underpowered, however, it will still be able to detect significant lipidomic differences between the *Gata6*-WT and -KO^{mye} pResMϕ, and importantly direct me to decide which lipids should be analysed using targeted MS/MS methods in order to validate my findings.

Lipids from GATA6-deficient cells were compared to those obtained from WT pResMϕ. The statistical analysis performed on the untargeted lipidomics data showed that the lipidome of *Gata6*-WT and -KO^{mye} pResMϕ was significantly different. 8% of putatively identified lipids increased significantly in *Gata6*-KO^{mye} pResMϕ. However, when looking at one of the most important lipid categories (FA), which comprises a diverse group of molecules, it was one of the lipid categories that changed the least (3%), and the changes in lipid levels were not consistent throughout the different lipid classes, thus suggesting that GATA6 has a regulatory effect on some, but not all, of the enzymes involved in FA metabolism.

The analysis of free fatty acids revealed a trend for longer chain unsaturated fatty acid levels to increase with GATA6 deficiency. Increases in unsaturated fatty acids has been shown to suppressed ABCA1 and ABCG1 transcripts and protein levels in human Mϕ[209]. However, I could find no evidence of transcriptional suppression of *Abca1* and *Abcg1* in *Gata6*-KO^{mye} pResMϕ, which could be due to different regulatory mechanisms found in humans and mice, or the fact that the increase in unsaturated fatty acid levels as not enough to elicit transcriptional changes.

Out of the 8 classes analysed, 2 showed significant changes in over 10 % of lipids, most of which were increased in *Gata6*-KO^{mye} pResMϕ. These 2 classes were SL and GPL, and they will be further discussed in the following sections.

5.3.1 SL

SL are located in the outer leaflet of the plasma membrane and interact with cholesterol-containing specific membrane domains known as 'lipid rafts'. These are small regions (10-200 nm) that play a key role in the compartmentalisation of cellular processes by stabilising protein-protein or protein-lipid interactions. SL are known for playing important roles in inflammation, activation of pro-inflammatory transcription factors in different cell types and induction of cyclooxygenase-2, leading to production of pro-inflammatory prostaglandins[210]. The presence of Caveolin-1 in a lipid raft creates an invagination of the membrane recognised as 'caveolae'. The data from section 5.2.2b) shows that the most significantly altered SL were long chain HexCer and SM, commonly found in the caveolar membrane fraction[171]. This could potentially contribute to a disturbance in caveolae leading to a less effective endocytosis. In the next Chapter I will validate the changes observed in SL using a targeted MS/MS approach.

5.3.2 GPL

Over 50% of all GPL were putatively identified as PC or PE. The host-defence process performed by Mϕ starts with the attachment, recognition and endocytosis of particles by the plasma membrane, the sealing vacuole and the digestion by the lysosome[211]. These events are accompanied by the secretion of various factors, and also by the induction of changes in the metabolism of membrane components, in particular GPL[212]. PG was the lipid category that presented the highest percentage of significantly increased GPL, with a trend for the shorter chain lipids to increase the most. Previous studies in alveolar Mϕ have shown that phosphatidylglycerol stimulates the phagocytosis by resMϕ but hardly affects the uptake by alveolar type II cells[213]. High concentrations of PG reduce the number of resMϕ in the alveolar space and deteriorate lung function.

However, further study is needed to clarify if PG has the same effects in resMϕ in the peritoneal cavity.

5.3.3 Conclusion

In summary, the data presented herein indicates that the transcription factor GATA6 might play an important role in the regulation of lipid metabolism in Mϕ causing an increase in lipid levels of different classes, however it is still unclear if this could be due to an increase in production or failure to degrade or utilize certain lipids. It is possible that the increase in lipid levels is caused by dysregulation of the lysosomes where lipids could accumulate. There are over 40 lysosome storage disorders known, including several sphindolipidoses where an intra-lysosomal accumulation of SL occurs[125]. Additionally, it has been shown that a marked decrease in Cathepsin E, a proteinase mainly localized on the endosomal compartment in Mϕ[168], can lead to a novel form of lysosome storage disorder[169]. Interestingly, the microarray data (Chapter 3) showed a downregulation of the cathepsin E gene (*Ctse*), indicating that cathepsin E could be involved in maintaining lipid homeostasis in Mϕ by preventing lipid degradation and consequently promoting an increase in lipid levels, such as the ones observed during the lipidomic analysis of *Gata6*-WT and -KO^{mye} ResMϕ.

Thus, it is imperative that a more in-depth analysis is performed, in order to validate the findings from this Chapter. Hence, in the next Chapter I will perform a targeted analysis of the sphingolipidome, using an LC-MS/MS analysis using lipid standards for the different SL classes. The advantage of this approach is that it will allow me to accurately identify and quantify individual lipid species, as opposed to using only accurate *m/z* and RT (untargeted approach).

Chapter 6 *Gata6* deficiency leads to sphingolipid accumulation in pResMϕ

6.1 Introduction

In Chapters 3 and 6, I presented transcriptomics and lipidomics data showing that GATA6 deficiency leads to lipid accumulation in pResMϕ, in particular SL. In order to further investigate this, here I will perform a targeted sphingolipidomic analysis of *Gata6*-WT and -KO^{mye} pResMϕ.

SL are a highly diverse family of bioactive lipids that contribute to numerous aspects of cell structure and signaling[118,127,214]. Numerous functions have been associated with all SL classes, such as sphingoid bases[215], Cer[215], SM[216], and both simple[217] and complex GSL, in Mϕ[218–220]. Furthermore, inhibition of *de novo* biosynthesis of SL, can modify Mϕ behaviour, *i.e.* cell cycle, apoptosis, autophagy and senescence [221–225].

The discovery and analysis of SL has benefited significantly from technological advances in lipid detection, analysis and quantitation using mass spectrometry. However, to date no method has been developed that can encompass all SL compounds in a structurally specific and quantitative manner[226]. The need for diverse methodologies arises from the fact that different SL classes dissociate into structurally different fragments corresponding to their sphingoid bases, N-acyl chains and polar headgroups[156,227]. Hence, a range of sample preparation, ionization modes and instrumental designs have been established to analyse particular SL classes individually by MS[123].

For further insight into the contribution of *Gata6* on SL metabolism, I established two different LC-MS/MS methods were using a quadrupole-linear-ion trap (6500 QTrap, operating in QQQ mode). These methods will allow me to quantify individual lipid species present in Mϕ. Unlike the global method I used in Chapter 6, which only provided unvalidated putative identifications. These MRM methods allow a far higher degree of selectivity and sensitivity, through the combination of chromatography and multiple reaction monitoring (MRM) detection.

Additionally, SL changes will be correlated with transcriptional changes observed in pResMϕ-KO^{mye}.

6.1.1 Aims

- To develop LC-MS/MS methods for targeted analysis of SL in pResMϕ.
- To perform a targeted analysis of the different classes of SL in *Gata6*-WT and -KO^{mye} pResMϕ.
- To assess how *Gata6* deficiency contributes to the Mϕ sphingolipidome.

6.2 Results

In this Chapter I will determine whether the myeloid deficiency of GATA6 is associated with significant changes in SL levels. An untargeted lipidomics workflow (described in Chapter 4) used to characterise the lipidomic differences between *Gata6*-KO^{mye} and –WT pResMϕ (Chapter 5) found that putatively identified SL were significantly increased in *Gata6*-KO^{mye}. Herein I will analyse *Gata6*-KO^{mye} and –WT pResMϕ samples using LC-MS/MS methods, in order to validate my findings. This approach will allow me to precisely investigate the contribution of GATA6 to the sphingolipidome of murine pResMϕ, and to determine which SL classes/species are the most affected by GATA6-deficiency.

To determine if the myeloid deficiency of GATA6 is associated with significant changes in SL levels, extracts from both cell types were analysed by LC-MS/MS. Briefly, peritoneal lavages collected from *Gata6*-WT and -KO^{mye} female mice were phenotyped and combined[7] (details in Materials and Methods, Chapter 2). pResMϕ were isolated by FACS (2 x 10^6 cells/replicate), and cell pellets stored at -80°C, until ready for extraction. When all three biological replicates were collected, lipids were extracted using a method optimised for SL recovery, previously described in the literature by R. L. Shaner *et al.*[156]. Next, in order to analyse SL, I needed to establish methods that would allow separation of different SL classes, since these were the lipids found to be different using the global lipidomics approach. The SL classes of interest were: sphingoid bases (So and Sa), Cer, SM, mono- (GlcCer and GalCer) and dihexocylCer (LacCer), Figure 6.1.

[7] Due to the low recovery of pResMϕ in GATA6 deficient mice, it was necessary to combine peritoneal lavages from ~15 mice. In order to obtain 2 x 10^6 cells/replicate.

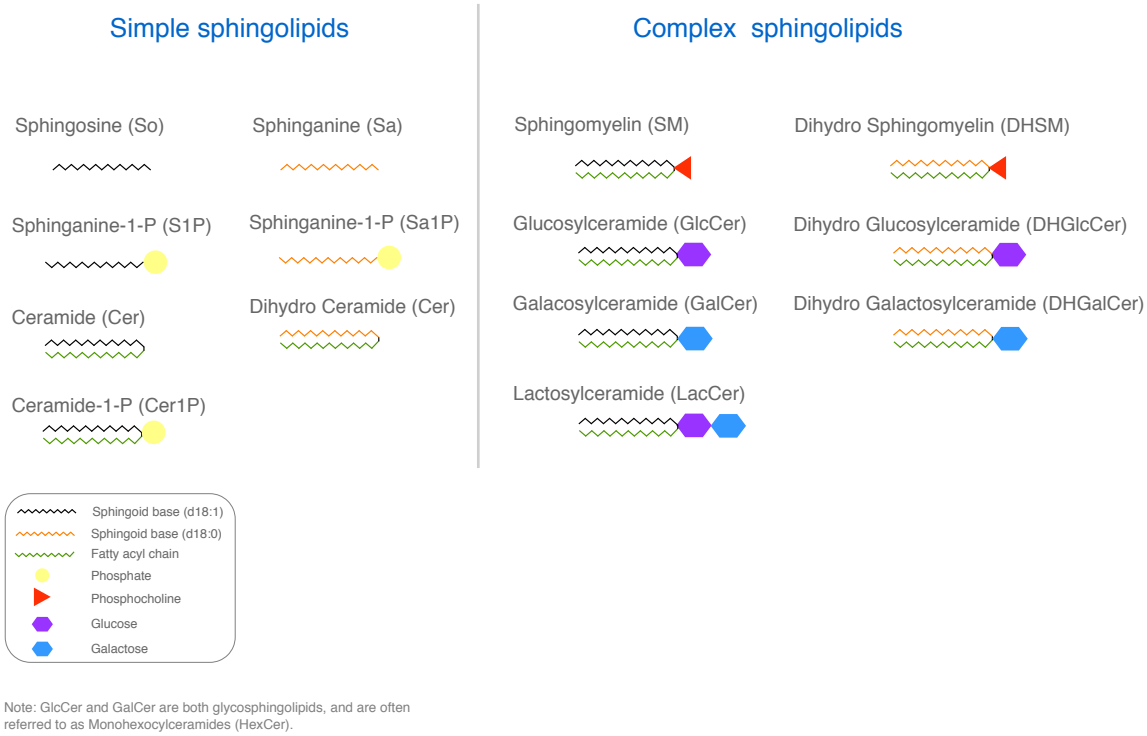


Figure 6.1 Higher order SL depicted symbolically to demonstrate complexity and compare similarities of added headgroups to the Cer base unit.

6.2.1 Establishment of LC-MS/MS methods for analysis of SL

The first, of the two approaches used in this Chapter, will allow the quantification of SL, and is be based on the LC method used for the untargeted lipidomics approach, described in Chapter 4, in combination with an MRM method, derived from the one described by R. L. Shaner *et al.*[156]. The second approach will complement the first, by separating GlcCer and galactosylCer, and is an adaptation of a method also described by R. L. Shaner *et al.*[156].

a) Quantitative analysis of SL in cell extracts by LC-MS/MS

The LC method I used for the targeted analysis of SL species was based on the LC conditions, described previously, for the untargeted analysis. Briefly, the lipid extracts were separated on a Accucore C18 column (150 x 2.1 mm, 2.6 µm) using an isocratic mobile phase A (H_2O:ACN 80:20 v/v) and a phase B (IPA: ACN, 70:30 v/v) at a flow rate of 0.4 mL/min over 35 min, with 4 mM ammonium acetate. Samples were maintained at 4 °C and the column at 25 °C. These LC conditions had already proven useful in the separation of SL (Chapter 5, untargeted lipidomics), thus adapting them for a targeted method allowed for the comparison of retention times between the two approaches.

The system used for these analyses was a Shimadzu LC-30 AD VP binary pump system coupled to a Shimadzu SIL-30 AC 200 autosampler coupled to a 6500 quadrupole linear-ion trap (QTrap) (AB SCIEX™, Warrington, UK) operating in a triple quadrupole mode.

Next, the optimal conditions for ionization and fragmentation were determined for each analyte. An internal standard mix developed by the LIPID MAPS Consortium was used, allowing the quantification of: sphingoid bases, as well as more complex species, such as SM, Cer, mono- and di-hexosylCer. The SL standard mix was dissolved at a concentration of 1 – 10 pmol/µL and infused into the ion source at a rate of 10 µL/min to determine the optimal ionization conditions by varying the declustering potential (DP) and entrance potential (EP) for the 6500 QTrap. After the Q1 settings were determined, product ion spectra were collected across a range of collision energies (CE), structurally specific

product ions were identified, and collision energies and collision cell exit potentials (CXP) were manipulated to produce optimal signal.

Once the settings were identified (Materials and Methods), the precursor/product ion pairs were placed in MRM methods as described in Materials and Methods. LC ESI-MS/MS elution profiles for the SL were confirmed (Figure 6.2).

i) Standard curves for quantitative analysis of SL by LC-MS/MS

Once the LC and MRM methods were established, standard curves were determined under these conditions. The Cer/Sphingoid Internal Standard Mixture II from Avanti was used, it contains So (C17 base), Sa (C17 base), Sphingosine-1-P (C17 base), Sphinganine-1-P (C17 base), Lactosyl(ß) C12 Cer, 12:0 SM, Glucosyl(ß) C12 Cer, 12:0 Cer, 12:0 Cer-1-P. The concentration of the individual components of the standard cocktail was 25 µM. These were serially diluted into the appropriate LC solvent (starting condition mobile phase) immediately before analysis to provide 0.5 – 1000 pmol of each standard per 20 µL injection. Each was then analysed by LC-MS/MS to generate the standard curves such as those shown in Figure 6.3, then to calculate the linear regression lines.

According to the literature, none of the lipid species present in the standard mix have been found in mammalian cells[228]. Additionally, it was previously reported that the signal response across the different chain lengths (C12:0 – C25:0) for all the SL classes is very similar, with little difference in cps due to the length of the fatty acyl chain, allowing the C12 standard to be used for all species[156].

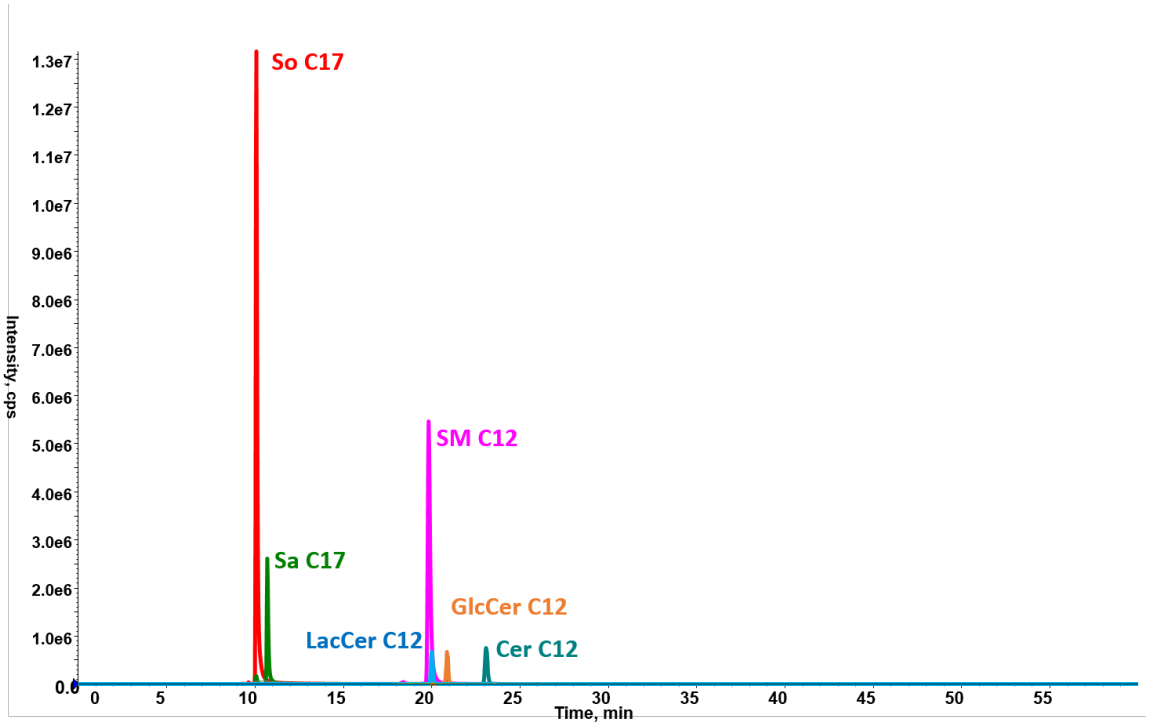


Figure 6.2 LC ESI-MS/MS elution profiles for the SL standards on reverse phase chromatography.

Shown are the elution profiles of sphingoid bases, Cer, LacCer, GlcCer and of an internal standard cocktail (500 pmol of each internal standard) from an Accucore C18 column (150 x 2.1 mm, 2.6 µm) and analysis by MRM in positive ionization mode as described under in the methods section (Chapter 2). The abbreviations identify the lipid classes: So C17, sphingosine d17:1; Sa C17, sphinganine d17:0; Cer C12, d18:1/C12:0 Cer; GlcCer C12, d18:1/C12:0 Glucosyl Cer; LacCer C12, d18:1/C12:0 Lactosyl Cer; SM C12, d18:1/C12:0 SM.

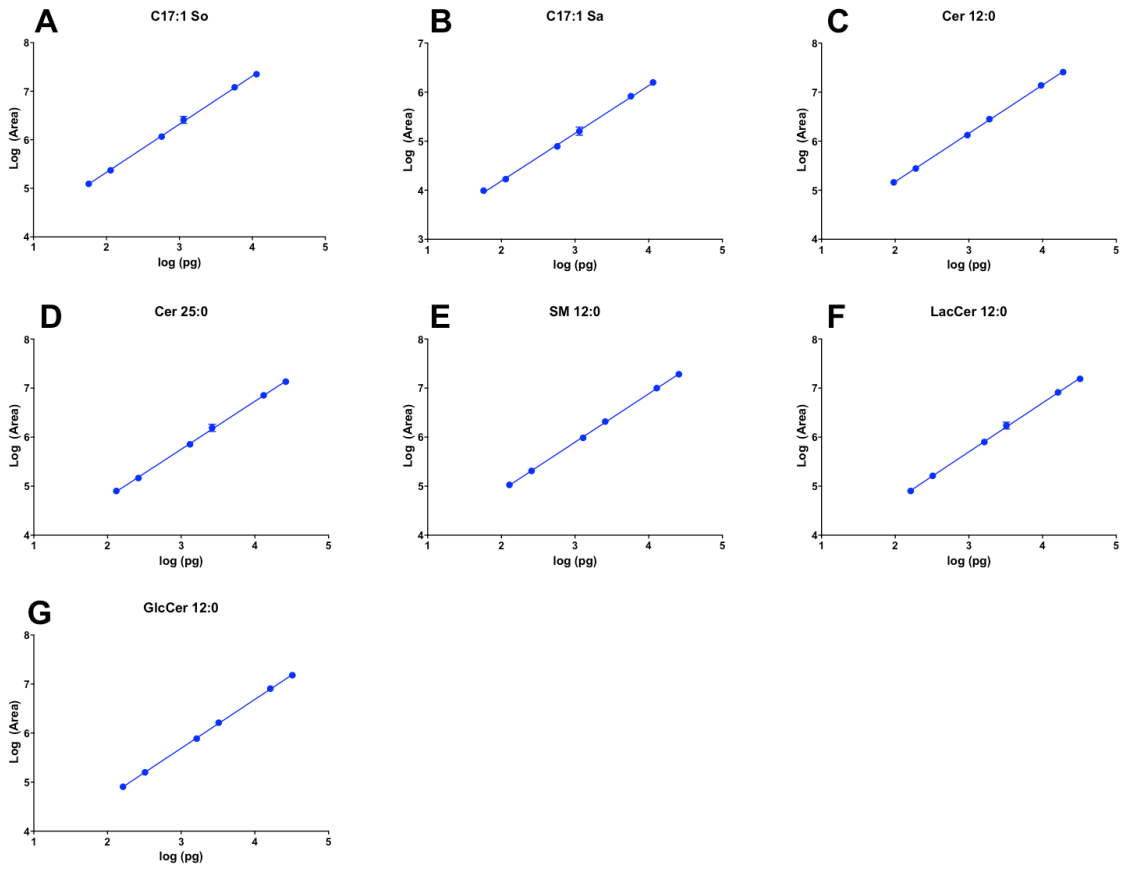


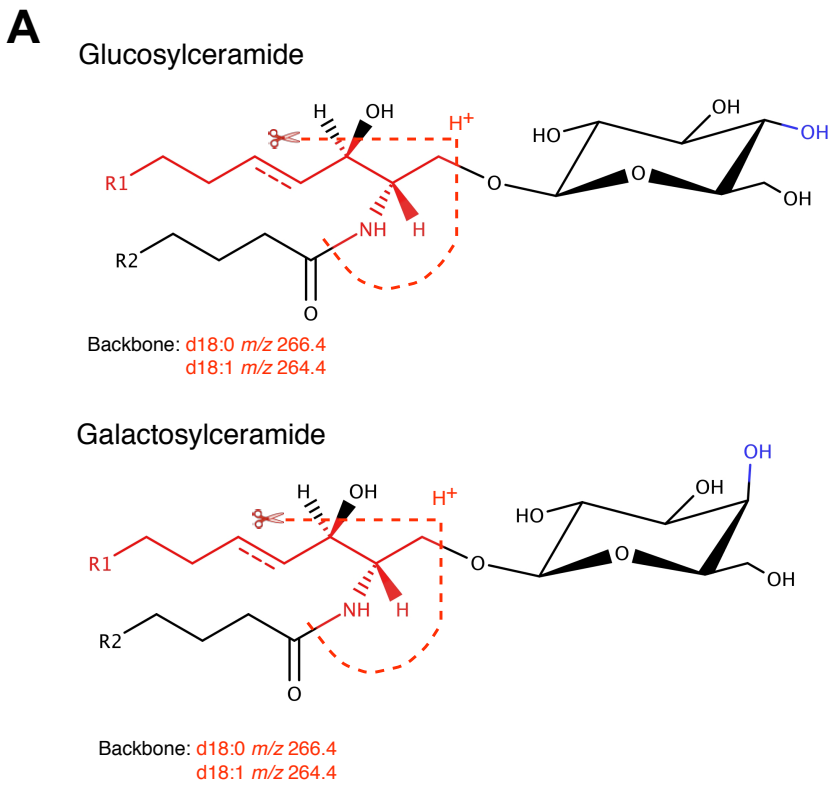
Figure 6.3 Signal response for SL standards using a 6500 Qtrap.

Each compound was analysed, in duplicate, over the shown range in amounts per injection using the LC-MS/MS protocol for sphingoid bases, Cer, SM, LacCer and GlcCer on reverse chromatography and positive ion mode MS/MS, using optimised ionisation, fragmentation and MRM conditions as described in Materials and Methods (Chapter 2).

b) MonohexosylCer (HexCer)

HexCer are Cer with one sugar residue joined in a β-glycosidic linkage at the 1-hydroxyl position, either glucose or galactose, which are epimers of each other. Thus, glucosyl- and galactosylCer have the same molecular formula, and fragmentation patern, but differ in the configuration of a stereocentre (depicted in blue, Figure 6.4A). The difference in configuration allows the separation of the two structures by LC (Figure 6.4B).

HexCer were analysed using the method described by R. L. Shaner *et al.*[156] (materials and methods), without any modifications, where it was established that standards GlcCer and GalCer with comparable backbones did not reveal any differences in ion yield and therefore, the C12-GlcCer internal standard is adequate for analysis of total HexCer (*i.e.*, where the stereochemistry of the carbohydrate is not specified).



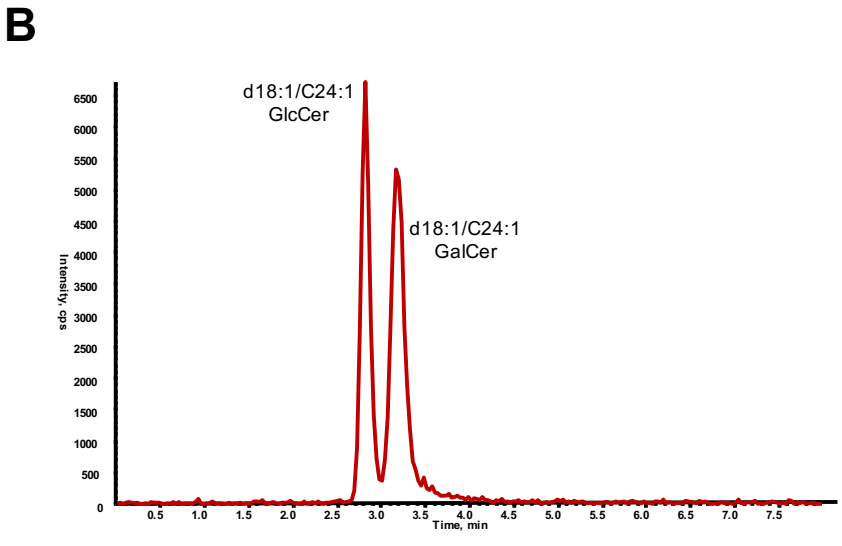


Figure 6.4 Separation of different monohexosylCer.

A- Depiction of the major fragmentations of different monohexosylCer (Glucosyl- and GalactosylCer). Depicted in the structures, are the monohexosyl groups at position 1, and R1 and R2 are the alkyl sidechains for the sphingoid base and fatty acid, respectively, which can vary in length and, to some extent, unsaturation. Red- major fragment obtained; Blue- are epimeric position where glusose and galactose differ. B- chromatographic separation of GlcCer and GalCer.

6.2.2 Analysis of SL in pResMϕ

The analysis of SL in pResMϕ extracts was conducted in MRM mode where specific precursor-product pairs are monitored during the LC elution. I compiled a list of SL putatively identified in Chapter 5, as well as, others known to be present in Mϕ[95,156]. The transitions for these lipid species were obtained from the literature, or derived from existing ones[156].

To quantify the amounts of the SL analytes of interest in *Gata6*-WT and -KO^{mye} pResMϕ extracts, the elution profiles for each MRM pair were examined. The areas under the peaks generated for both analytes and internal standards were integrated with the software Multiquant[97]. I first established SL levels in *Gata6*-WT pResMϕ and then compared them with those in *Gata6*-KO^{mye} pResMϕ. The analysis of various SL classes is described in detail in the following sections.

a) Sphingoid bases

Here I analysed long-chain bases (So, Sa) and their derivatives (sphingoid base 1-phosphates, glycosphingosine), described in Table 6.1 (Figure 6.5 shows a representative example of LC-MS/MS analysis of So and Sa in pResMϕ cells), they are highly bioactive compounds that appear as intermediates of SL metabolism and cell signalling, and are elevated on disruption of SL metabolism in disease[229]. I first established the lipid levels in *Gata6*-WT pResMϕ and was able to quantify 4 of the 5 lipid species in Table 6.1. Additionally, I found that the relative amounts of sphingoid bases varies significantly, and that So were more abundant than Sa (Figure 6.6C and D, respectively).

Next, I compared *Gata6*-WT and -KO^{mye} pResMϕ and found the following changes. Of the four lipids analysed, in this class, only the levels of So were significantly changed in *Gata6*-KO^{mye} (Figure 6.6A), displaying a 30 % increase. Both Sa and So are precursors for DHSL and SL (Figure 6.7), which could potentially lead to higher levels of SL vs DHSL when comparing *Gata6*-KO^{mye} with -WT pResMϕ. In particular DH-Cer and Cer, the first lipid species produced directly from Sa and So, which will be analysed in the next section.

Table 6.1 Parent to daughter ion transitions used for measuring sphingoid bases, and concentrations found in *Gata6*-WT pResMϕ. ND, not detected.

Lipid	Q1 Mass (Da)	Q3 Mass (Da)	Concentration (pmol/10^6 cells)
C17:0 Sa	288.44	270.2	Standard
C17:0 Sa-1-P	368.4	252.2	Standard
C17:1 So	286.45	268.2	Standard
C17:1 So-1-P	366.3	250.2	Standard
C18:0 Sa	302.4	252.2	0.2
C18:0 Sa-1-P	382.4	266.4	ND
C18:1 So	300.4	282.4	6.1
C18:1 So-1-P	380.4	264.4	5.0
Lyso-GlcCer	490.4	252.2	0.4

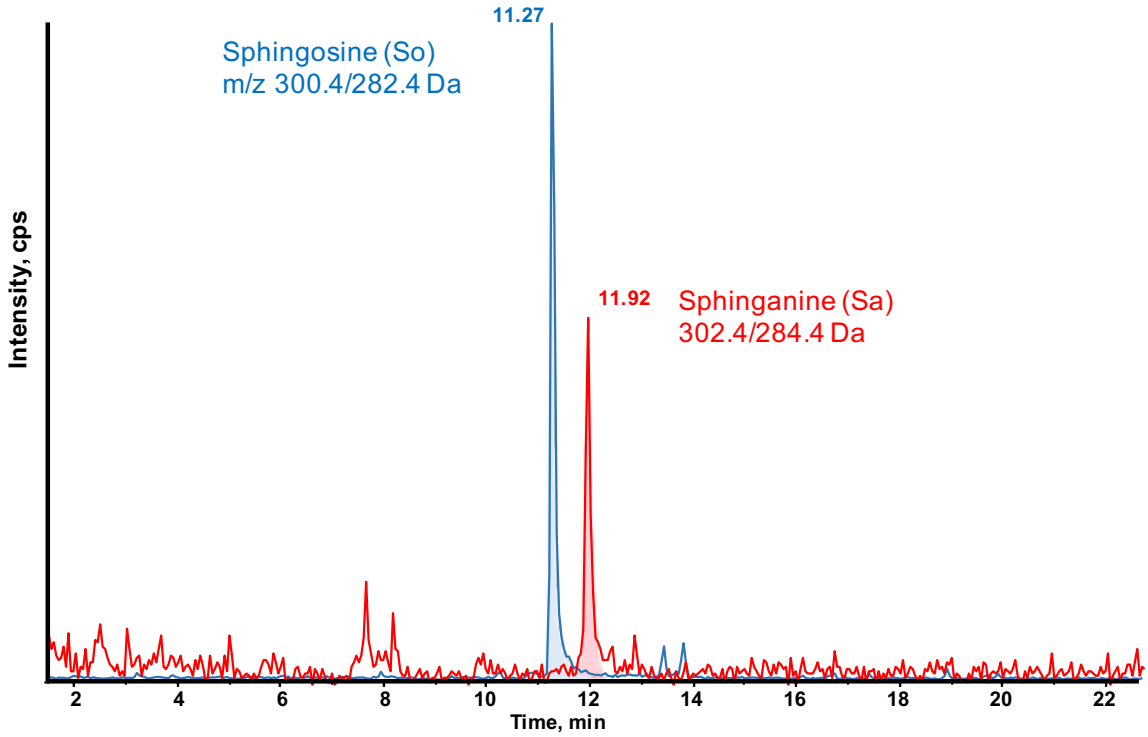


Figure 6.5 Representative example of the LC-MS/MS analysis of sphingoid bases in pResMϕ cells.

Lipid extracts from pResMϕ cells were analysed by LC-MS/MS. Shown are the chromatograms obtained for So (blue) and Sa (Red), and the respective integrated areas (shaded area).

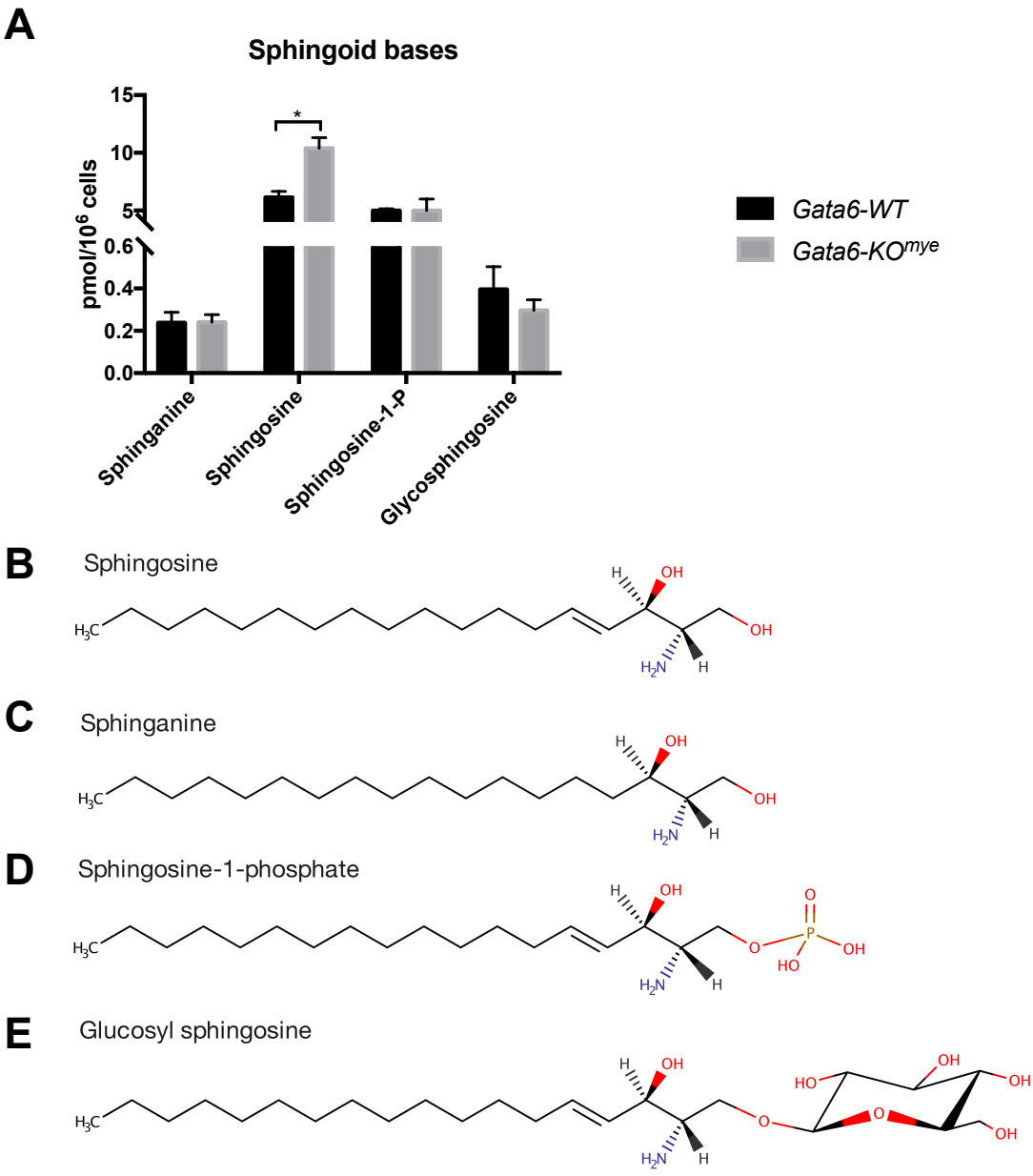


Figure 6.6 Quantitative analysis of sphingoid bases in *Gata6*-KO^{mye} and -WT pResMϕ cells.

The amounts of these lipids were measured by LC-MS/MS using the internal standard cocktail as described in Materials and Methods (Chapter 2). Shown are the means ± SEM for three separate replicates with cells pooled from 15 *Gata6*-KO^{mye} and 15 -WT animals. A - shows the amounts of sphingoid bases; B - E, chemical structures of a So, Sa, S1P and glucosyl sphingosine, respectively.

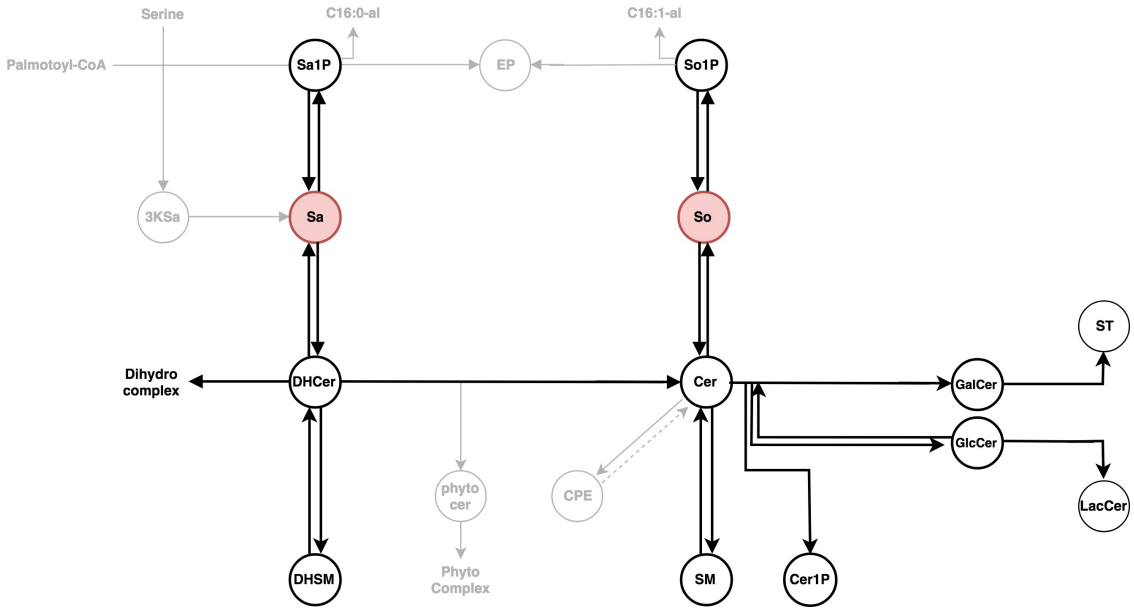


Figure 6.7 Steps of SL biosynthesis.

The figure depicts metabolites (circles) that participate in early steps of SL biosynthesis: Sa1P, sphinganine-1-phosphate; EP, ethanolamine phosphate; S1P, sphingosine-1-phosphate; 3KSa, 3-ketosphinganine; Sa, Sa; So, sphingosine; DH-Cer, dihydroCer; Cer, Cer; DH-SM, dihydrosphingomyelin; SM, sphingomyelin; GalCer, galactosylCer; GlcCer, GlcCer; Cer1P, Cer-1-phosphate; LacCer, lactosylCer. The metabolites analysed by LC-MS, as previously described, are highlighted in black. Sphinganine (Sa) and sphingosine (So), are depicted in red.

b) Cer

Here I analysed Cer and CerP species, described in Table 6.2 (Figure 6.8 shows a representative example of LC-MS/MS analysis of Cer and CerP in pResMϕ cells), they are considered backbone signalling subspecies of SL[134]. I first established the lipid levels in *Gata6*-WT pResMϕ and found that I could not detect any of the DH-Cer species, Table 6.2. Regarding the analysis of Cer1P, only the C16 and C24 species were observed, namely d18:0/24:0-CerP, d18:0/24:1-CerP, d18:1/16:0-CerP, d18:1/24:0-CerP. The relative amounts of Cer varied considerably for different chain-lengths (Figure 6.9B), with C16:0 being most abundant, followed by C24:1 and C24:0. Comparing saturated with monounsaturated forms, revealed that the latter presented lower levels for all species, except C24.

Next, I compared the levels of Cer and CerP in *Gata6*-KO^{mye} and -WT pResMϕ, and observed that even though all lipid species tended to be increased in *Gata6*-KO cells, d18:0/16:0-CerP was the only CerP to be significantly different (Figure 6.9). With respect to Cer, 2 species were increased significantly upon *Gata6* knockdown, d18:1/24:0-Cer and d18:1/25:0-Cer.

Table 6.2 Parent to daughter ion transitions used for measuring Cer and CerP, and concentrations found in *Gata6*-WT pResMϕ. ND, not detected.

	N-Acyl	Q1 Mass (Da)	Q3 Mass (Da)	Concentration (pmol/10^6 cells)
DHCer	C16:0	540.7	266.4	ND
	C18:0	568.7	266.4	ND
	C20:0	596.6	266.4	ND
	C22:0	624.6	266.4	ND
	C24:0	652.9	266.4	ND
	C24:1	650.9	266.4	ND
	C26:0	680.7	266.4	ND
	C26:1	678.7	266.4	ND
DH-CerP	C24:0	732.9	266.4	0.1
	C24:1	650.6	266.4	1.0
CerP	C12:0	562.5	264.5	Standard
	C16:0	618.7	264.4	131.8
	C24:0	730.9	264.4	3.2
Cer	C12:0	482.6	264.2	Standard
	C16:0	538.7	264.4	90.5
	C18:0	566.5	264.4	2.9
	C18:1	564.5	264.4	0.2
	C20:0	594.6	264.4	2.4
	C20:1	592.6	264.4	0.3
	C20:0	622.6	264.4	4.7
	C22:1	620.6	264.4	1.2
	C24:0	650.6	264.4	3.4
	C24:1	648.6	264.4	10.9
	C25:0	664.8	264.4	0.6
	C26:0	678.7	264.4	ND
	C26:1	676.7	264.4	0.4

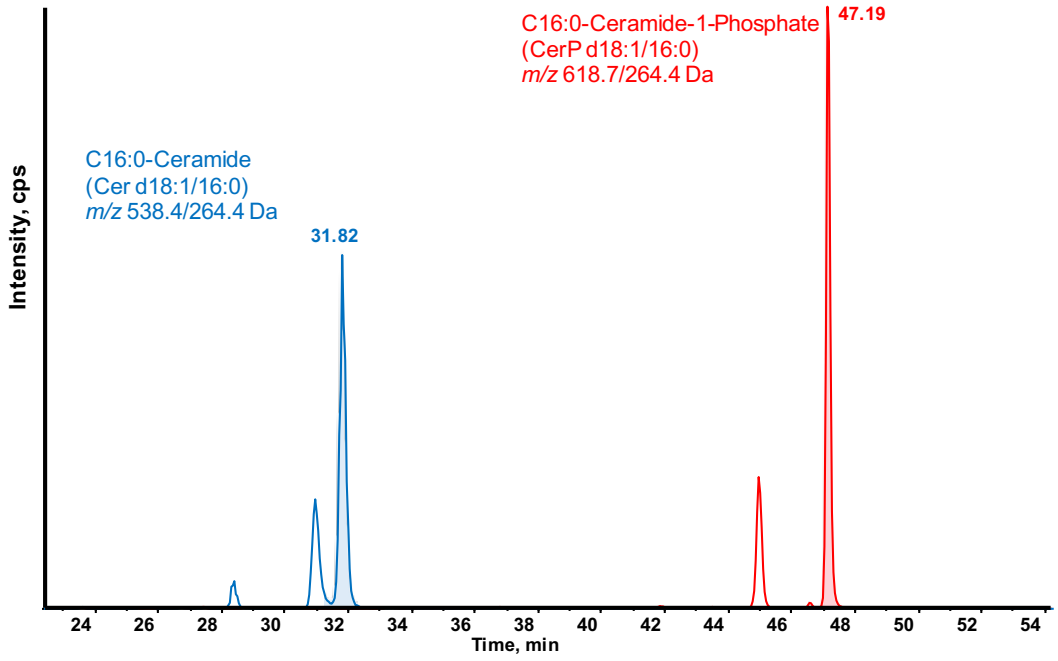


Figure 6.8 Representative example of the LC-MS/MS analysis of Cer in pResMϕ cells.
Lipid extracts from pResMϕ cells were analysed by LC-MS/MS. Shown are the chromatograms obtained for Cer d181/16:1 (blue) and CerP d18:1/16:0 (Red), and the respective integrated areas (shaded area).

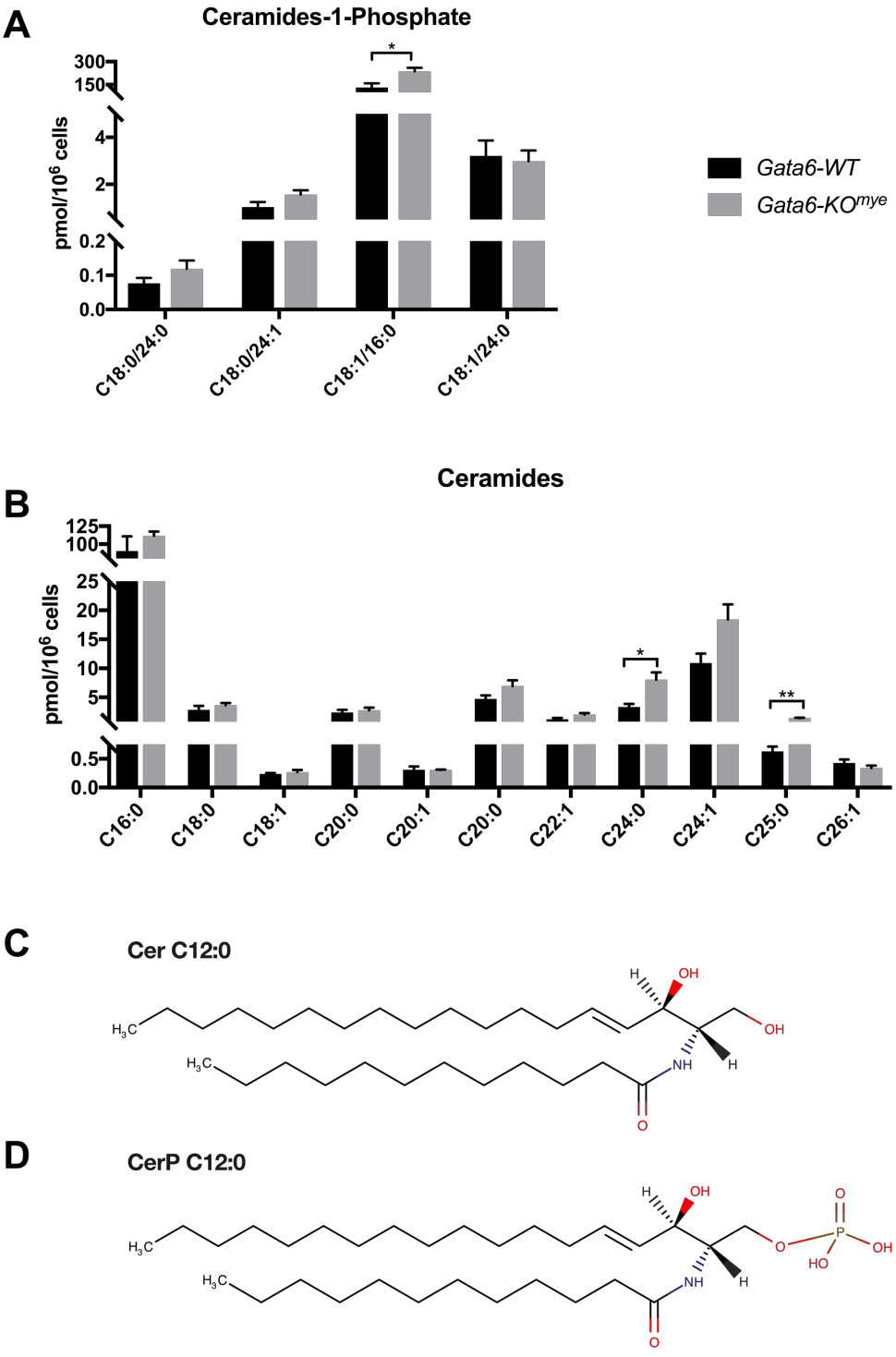


Figure 6.9 Quantitative analysis of Cer and Cer-1-phosphate in *Gata6*-KO^{mye} and -WT pResMϕ cells.

The amounts of these lipids were measured by LC-MS/MS using the internal standard cocktail as described in Materials and Methods (Chapter 2). Shown are the means ± SEM for three separate replicates with cells pooled from 15 *Gata6*-KO^{mye} and 15 -WT animals. A - shows the amounts of Cer-1-phosphates; B - shows the amounts of C and D, chemical structures of Cer d18:1/12:0 and Cer-1-phosphate d18:1/12:0, respectively.

c) HexCer

Here I analysed DH-HexCer and Cer species, described in Table 6.3. Figure 6.10 shows a representative example of LC-MS/MS analysis of Cer and CerP in pResMϕ cells. I first established the lipid levels in *Gata6*-WT pResMϕ and found that whilst all the HexCer species were detected, only 50 % of DH-HexCer were above the LOD. Furthermore, the levels of dihydro forms are ca. 45 times lower than those of HexCer for the same chain-lengths Table 6.3.

Regarding the comparison of DH-HexCer/HexCer, the ratio between these 2 classes increases from 45 (*Gata6*-WT pResMϕ) to 55, when comparing amounts in *Gata6*-KO^{mye} pResMϕ (Figure 6.11A and B). Analysis of DH-HexCer (Figure 6.11D) showed that 5 out of 7 detected lipids increased significantly in *Gata6*-KO^{mye} pResMϕ (Figure 6.11A), which is considerably higher than what was found for HexCer, where only 7 out of 12 molecules were significantly changed.

Many cells do not contain galacto-family SL, so distinguising between GlcCer and GalCer is not required[156]. Nonetheless, there is the possibility that GalCer was enriched in *Gata6*-KO^{mye} pResMϕ, therefore, it was necessary to separate GlcCer and GalCer. SL analysis using the method described in section 6.2.1b) showed no evidence of GalCer lipids in either pResMϕ from -KO^{mye} and -WT (Figure 6.12).

Table 6.3 Parent to daughter ion transitions used for measuring HexCer and concentrations found in *Gata6*-WT pResMϕ. ND, not detected.

	N-Acyl	Q1 Mass (Da)	Q3 Mass (Da)	Concentration (pmol/10^6 cells)
DH-HexCer	C14:0	674.5	266.4	ND
	C14:1	672.5	266.4	ND
	C16:0	702.6	266.4	0.2
	C16:1	700.6	266.4	0.0
	C18:0	730.6	266.4	ND
	C18:1	728.6	266.4	0.3
	C20:0	758.6	266.4	0.3
	C20:1	756.6	266.4	ND
	C22:0	786.7	266.4	0.2
	C22:1	784.6	266.4	ND
	C24:0	814.9	266.4	0.2
	C24:1	812.7	266.4	0.3
	C26:0	842.7	266.4	ND
	C26:1	840.7	266.4	ND
HexCer	C12:0	644.7	264.5	Standard
	C14:0	672.5	264.4	0.8
	C16:0	700.6	264.4	11.2
	C16:1	698.5	264.4	1.0
	C18:0	728.6	264.4	0.7
	C18:1	726.6	264.4	16.1
	C20:0	756.6	264.4	1.4
	C20:1	754.6	264.4	0.6
	C22:0	784.6	264.4	5.7
	C22:1	782.6	264.4	1.6
	C24:0	812.7	264.4	7.1
	C24:1	810.7	264.4	11.5
	C26:0	840.7	264.4	2.5
	C26:1	838.7	264.4	2.5

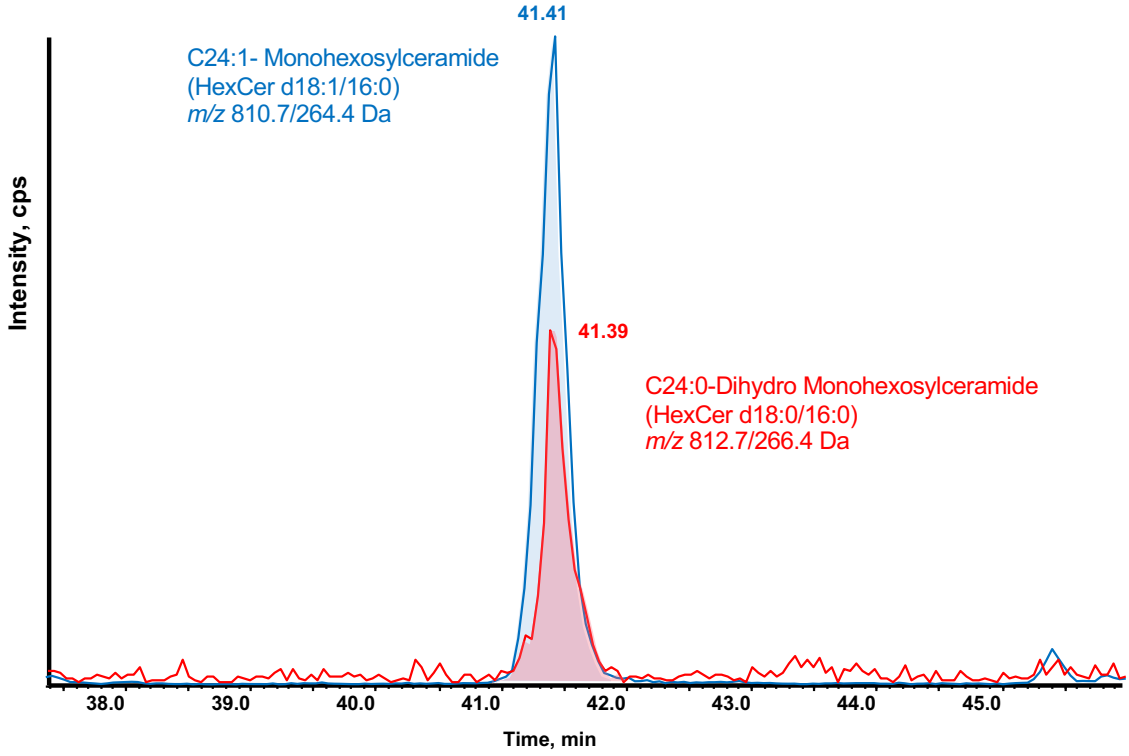


Figure 6.10 Representative example of the LC-MS/MS analysis of HexCer in pResMϕ cells.

Lipid extracts from pResMϕ cells were analysed by LC-MS/MS. Shown are the chromatograms obtained for HexCer d18:1/16:0 (blue) and HexCer d18:0/16:0 (Red), and the respective integrated areas (shaded area).

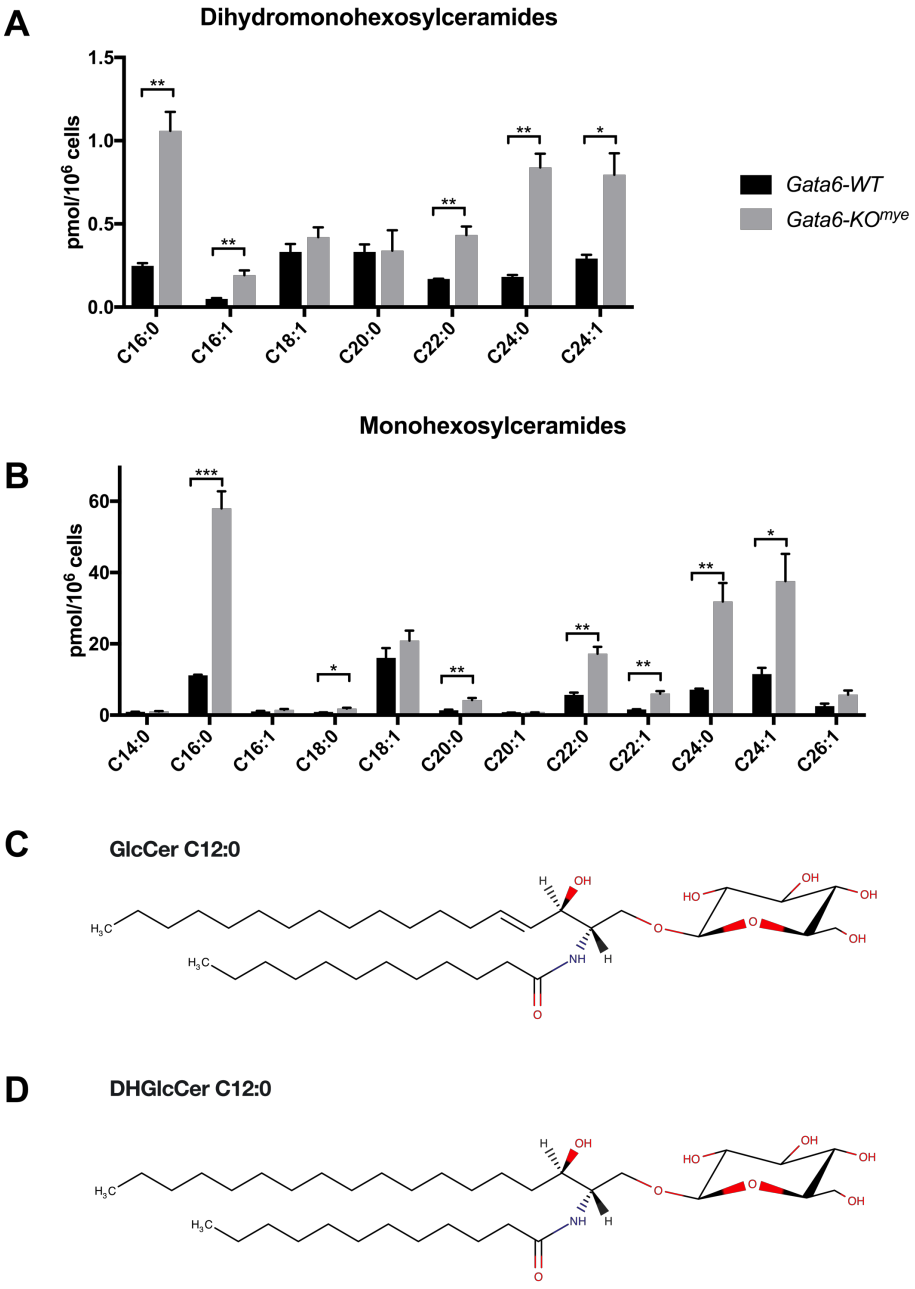


Figure 6.11 Quantitative analysis of DH HexCer and HexCer in *Gata6*-KO^{mye} and -WT pResMϕ cells.

The amounts of these lipids were measured by LC-MS/MS using the internal standard cocktail as described in Materials and Methods (Chapter 2). Shown are the means ± SEM for three separate replicates with cells pooled from 15 *Gata6*-KO^{mye} and 15 -WT animals. A - shows the amounts of DH HexCer; B - shows the amounts of HexCer. C and D, chemical structures of an HexCer (GlcCer d18:1/12:0) and an DH HexCer (DH-GlcCer d18:1/12:0), respectively.

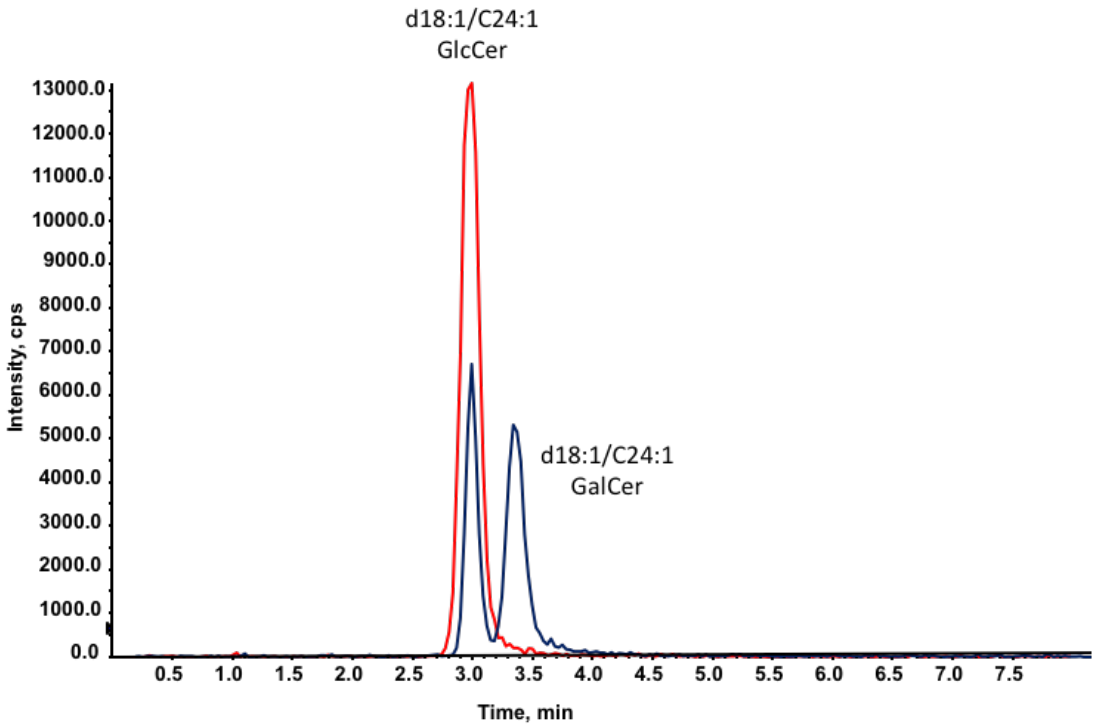


Figure 6.12 *Gata6*-KO^{mye} pResMϕ show no evidence of GalCer lipids.

LC ESI-MS/MS elution profiles for HexCer on normal phase chromatography. The figure shows a representative chromatogram of the separation of GluCer and GalCer (d18:1/C24:1-GlcCer and -GalCer) using a Supelco 2.1mm i.d. × 25 cm LC-Si column and the conditions described under "Materials and Methods.". Blue- standards; Red- *Gata6*-KO^{mye} pResMϕ cells.

d) LactosylCer

Here, I analysed the LacCer species, described in Table 6.4. Figure 6.13 shows a representative example of LC-MS/MS analysis of LacCer in pResMϕ cells. I first established their levels in *Gata6*-WT pResMϕ and found that out of the 11 LacCer species analysed, 4 were below the limit of detection.

Next, I compared LacCer levels in *Gata6*-WT and -KO^{mye}, and observed that there was a noticeable increase, in all LacCer species, in *Gata6*-KO^{mye} pResMϕ, however, only the changes in d18:1/18:0-LacCer, d18:1/22:0-LacCer, d18:1/24:0-LacCer, d18:1/24:1-LacCer were statistically significant (Figure 6.14).

Table 6.4 Parent to daughter ion transitions used for measuring LacCer and concentrations found in *Gata6*-WT pResMϕ. ND, not detected.

	N-Acyl	Q1 Mass (Da)	Q3 Mass (Da)	Concentration (pmol/10^6 cells)
LacCer	C12:0	806.7	264.4	Standard
	C16:0	862.7	264.4	0.2
	C18:0	890.7	264.4	0.1
	C18:1	888.6	264.4	ND
	C20:0	918.7	264.4	0.4
	C20:1	816.7	264.4	ND
	C22:0	946.8	264.4	0.2
	C22:1	844.7	264.4	0.1
	C24:0	974.8	264.4	0.2
	C24:1	972.7	264.4	0.2
	C26:0	1002.8	264.4	ND
	C26:1	1000.8	264.4	ND

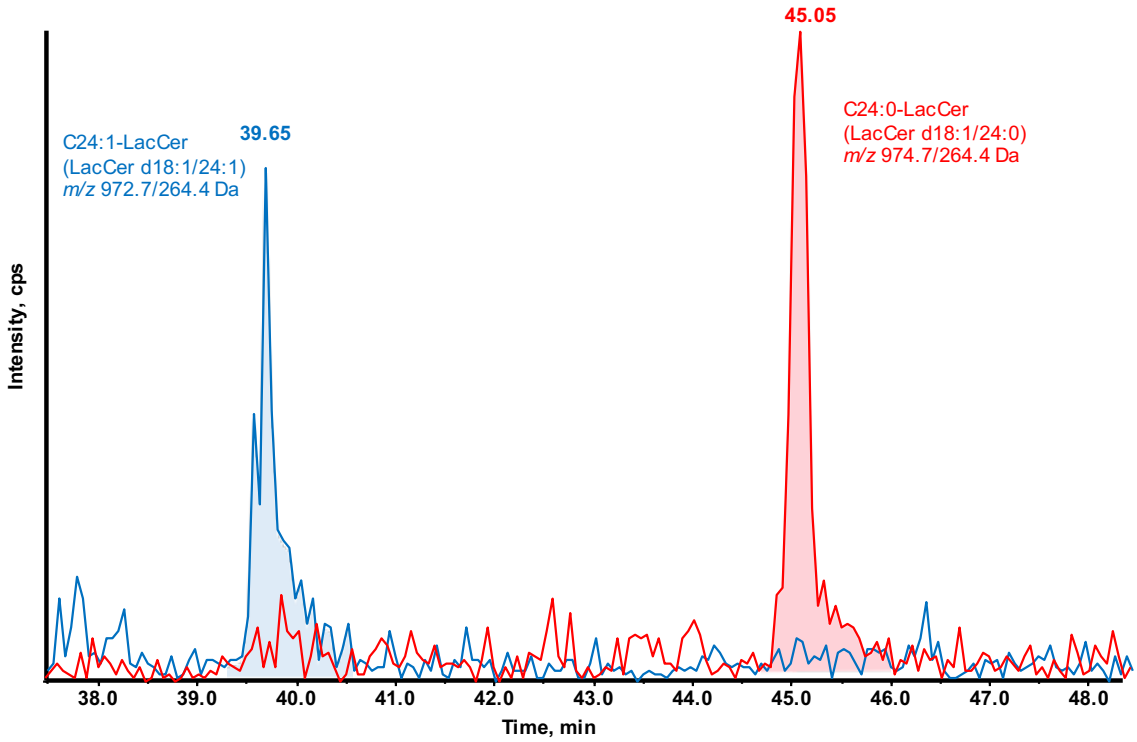


Figure 6.13 Representative example of the LC-MS/MS analysis of LacCer in pResMϕ cells.

Lipid extracts from pResMϕ cells were analysed by LC-MS/MS. Shown are the chromatograms obtained for LacCer d18:1/24:1 (blue) and LacCer d18:1/24:0 (Red), and the respective integrated areas (shaded area).

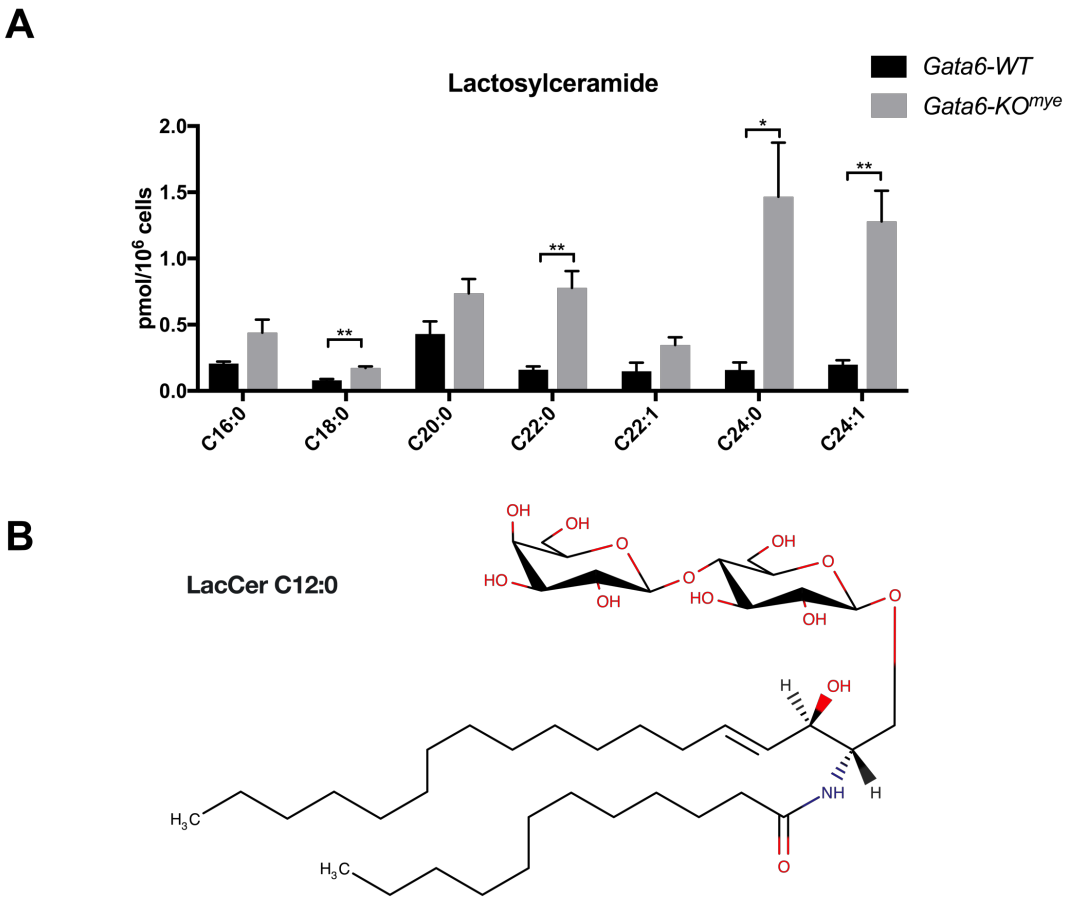


Figure 6.14 Quantitative analysis of LacCer in *Gata6*-KO^{mye} and -WT pResMϕ cells.

The amounts of these lipids were measured by LC-MS/MS using the internal standard cocktail as described in Materials and Methods (Chapter 2). Shown are the means ± SEM for three separate replicates with cells pooled from 15 *Gata6*-KO^{mye} and 15 -WT animals. A - shows the amounts of LacCer; B structure of LacCer d18:1/12:0.

e) Sphingomyelin

I analysed DH-SM and SM species, described in Table 6.5. Figure 6.15 shows a representative example of LC-MS/MS analysis of DH-SM and SM in pResMϕ cells. I first established the lipid levels in *Gata6*-WT pResMϕ and found that I could not detect the C14 DH-SM species and that C16 was the most abundant of the 7 species analysed, Table 6.5. Notably, the relative amounts of SM varied considerably for different chain-lengths (Figure 6.16B), with C16:0 being most abundant, followed by C24:1 and C24:0. Comparing saturated with monounsaturated forms revealed that the latter were present at lower levels for all species, except C24.

Next, I compared the levels of DH-SM and SM in *Gata6*-KO^{mye} and -WT pResMϕ. All lipid species tended to be increased in *Gata6*-KO cells, however, of the 6 DH-SM analysed, only 4 were significantly altered. With respect to SM, out of the 13 SM detected, 8 lipids were significantly increased, representing roughly half of all SM species analysed. A similar finding is seen DH-SM, *i.e*. 65% of lipids are changed, which is higher than that of their precursors DH-Cer and Cer (roughly 25%).

Table 6.5 Parent to daughter ion transitions used for measuring SM and concentrations found in *Gata6*-WT pResMϕ. ND, not detected.

	N-Acyl	Q1 Mass (Da)	Q3 Mass (Da)	Concentration (pmol/10^6 cells)
DH-SM	C14:0	677.6	184.4	ND
	C16:0	705.6	184.4	44.2
	C18:0	733.8	184.4	1.9
	C20:0	761.6	184.4	1.1
	C22:0	789.7	184.4	3.2
	C24:0	817.9	184.4	5.5
	C26:0	845.7	184.4	0.0
SM	C12:0	647.7	184.1	Standard
	C14:0	675.5	184.4	1.4
	C16:0	703.6	184.4	531.5
	C16:1	701.6	184.4	32.4
	C18:0	731.8	184.4	22.4
	C18:1	729.6	184.4	8.4
	C20:0	759.6	184.4	11.2
	C20:1	757.6	184.4	2.0
	C22:0	787.6	184.4	25.6
	C22:1	785.6	184.4	4.3
	C24:0	815.9	184.4	49.8
	C24:1	813.7	184.4	125.5
	C26:0	843.7	184.4	0.1
	C26:1	841.7	184.4	0.5

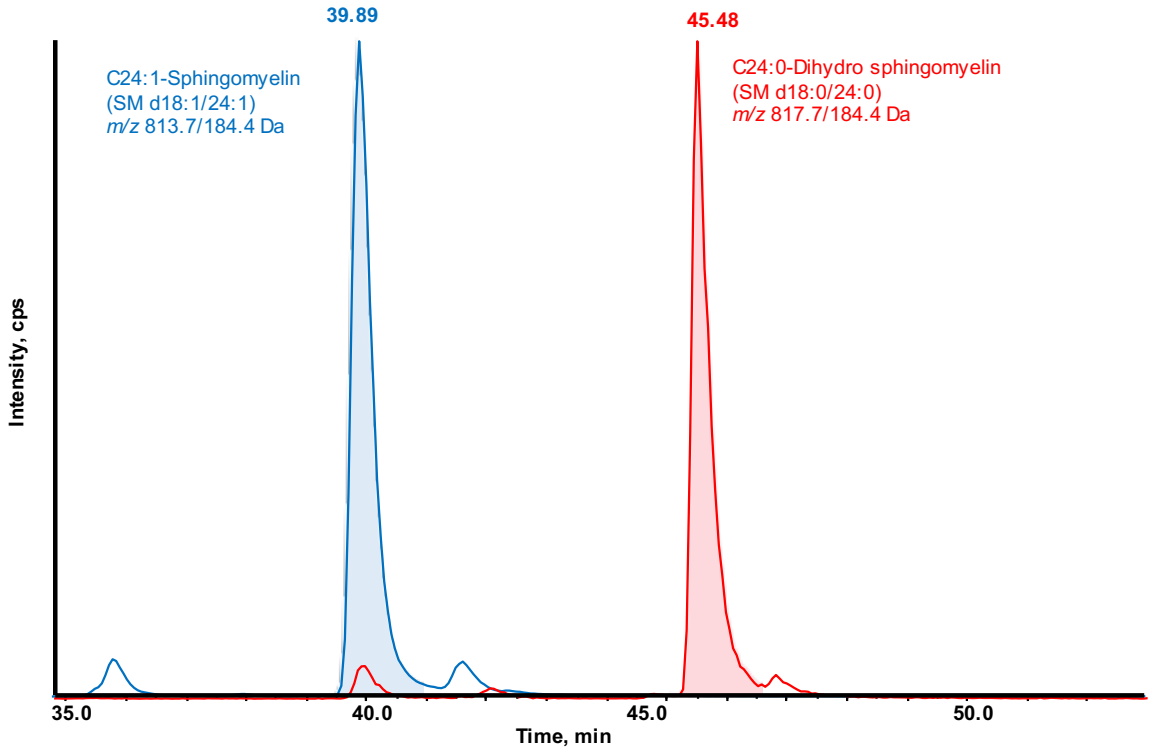


Figure 6.15 Representative example of the LC-MS/MS analysis of SM in pResMϕ cells.
Lipid extracts from pResMϕ cells were analysed by LC-MS/MS. Shown are the chromatograms obtained for SM d181/24:1 (blue) and SM d18:0/24:0 (Red), and the respective integrated areas (shaded area).

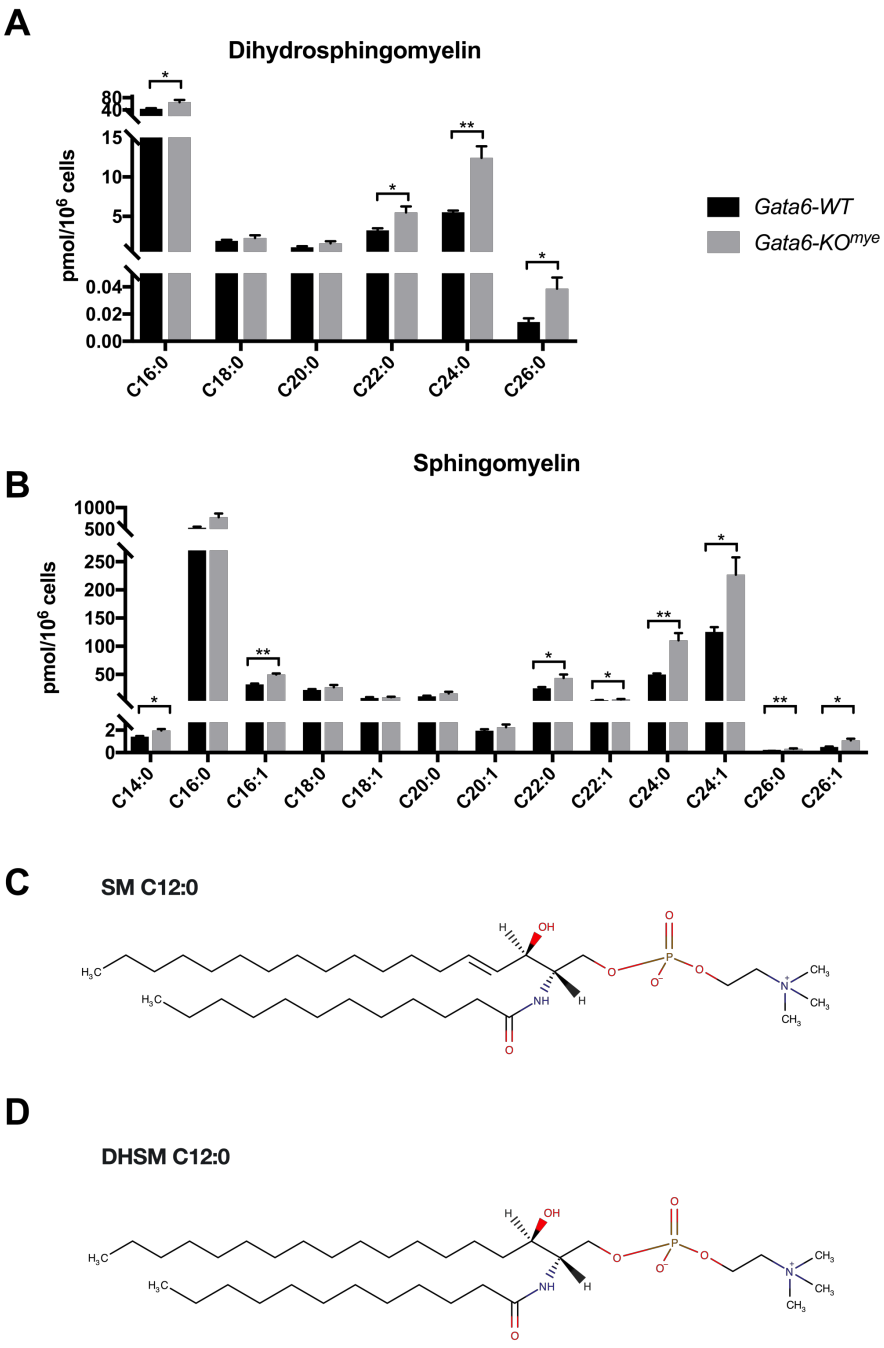


Figure 6.16 Quantitative analysis of DH SM and SM in *Gata6*-KO^{mye} and -WT pResMϕ cells.

The amounts of these lipids were measured by LC-MS/MS using the internal standard cocktail as described in Materials and Methods (Chapter 2). Shown are the means ± SEM for three separate replicates with cells pooled from 15 *Gata6*-KO^{mye} and 15 -WT animals. A - shows the amounts of DH SM; B - shows the amounts of SM. C and D, chemical structures of an SM d18:1/12:0 and an DH SM d18:1/12:0, respectively.

6.2.3 GATA6 deficiency leads to a significant increase in SL and DHSL in pResMϕ

Figure 6.17 summarises all of the subspecies of SL that have been analysed in *Gata6*-KO^{mye} and -WT pResMϕ. Volcano plots were generated, representing fold-change and p-value of lipids in *Gata6*-KO^{mye} relative to -WT pResMϕ cells, clearly show that GATA6 deficiency leads to significant increases in SL levels, where the majority of altered lipids have long and saturated fatty acyl chains.

A comparison between the number of lipids significantly changed for each class, revealed that the most affected lipid class by GATA6 deficiency was DH-HexCer, followed by DH-SM and SM. Considering that DH-Cer/Cer are the precursors for DH-SM/SM, the fact that the changes observed in DH-Cer/Cer do not translate to DH-SM/SM, suggests that the enzymes involved in the metabolism of DH-SM/SM might be affected by GATA6 deficiency, or that the assay might be close to its limits of detection. Likewise, DH-HexCer/HexCer increase, in *Gata6*-KO^{mye}, cannot be explained by variations in substrate (DH-Cer/HexCer).

In the next section I will explore the possibility that the dysregulation of SL metabolism related genes directly impacts SL levels.

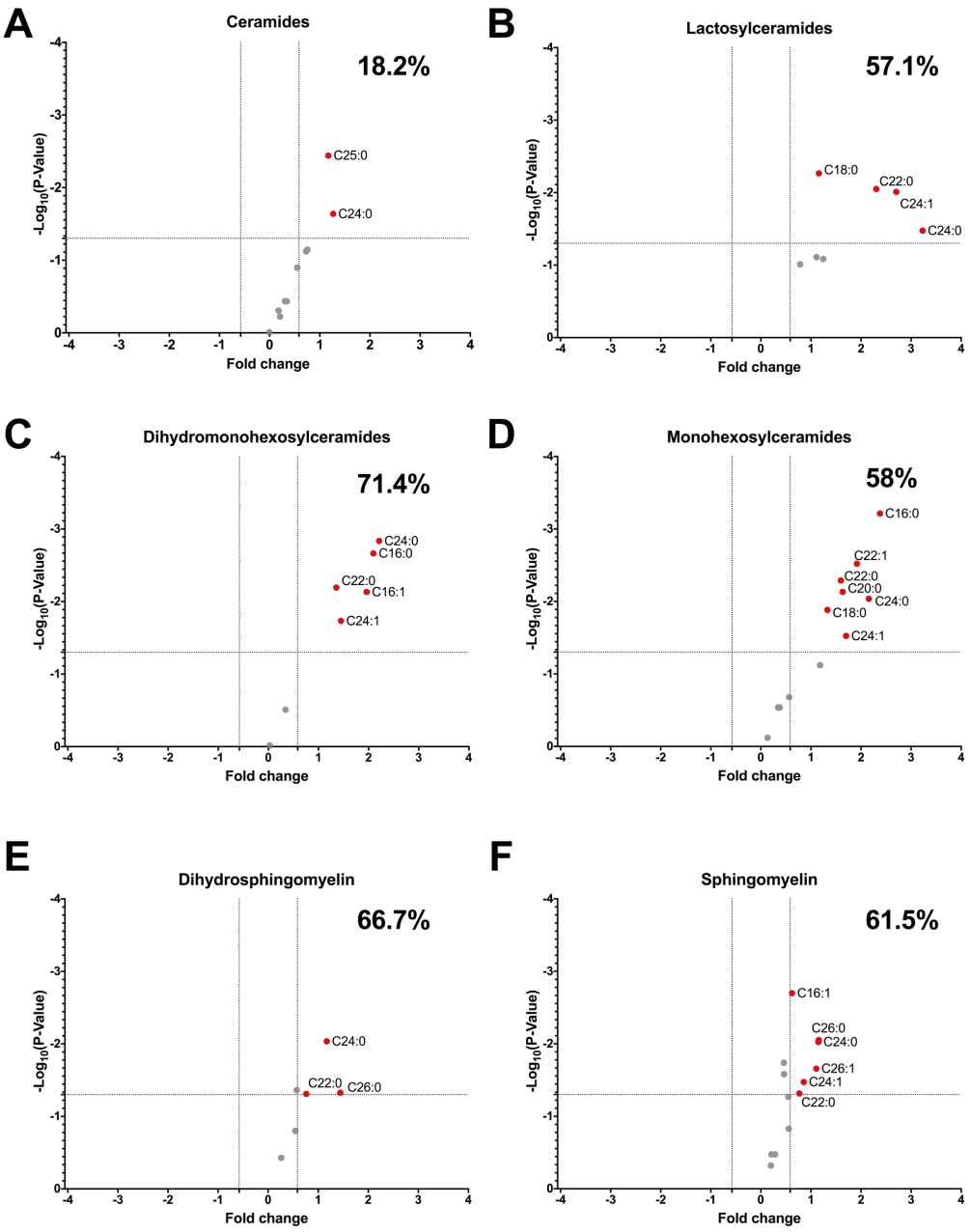


Figure 6.17 GATA6 deficiency leads to a significant increase in SL and DH-SL in pResMϕ.
Lipid extracts from *Gata6*-KO^{mye} and -WT pResMϕ cells. SL and DHSL were quantified by LC-MS/MS on a 6500 QTrap mass spectrometer (AB Sciex). Volcano plots representing fold-change and p-value of lipids in *Gata6*-KO^{mye} relative to -WT pResMϕ cells. Statistical significance determined without correction for multiple comparisons, with alpha=0.05%. Each row was analyzed individually, without assuming a consistent SD. Differential levels were classified as significant when P<0.05 and FC>1.5 (grey lines). Numbers on the upper right corner of each plot correspond to the percentage of significantly change lipids in each category.

6.2.4 Correlation between lipidomics and transcriptomics data reveals a potential mechanism for SM accumulation

The main changes in lipid levels observed in *Gata6*-KO^{mye} pResMϕ were the increase of DH-SM/SM and DH-GlcCer/GlcCer, however these could not be explained by substrate availability, given that only 18.2 % of Cer were significantly altered. Thus, I next analysed the expression of genes involved in SL metabolic pathways. Figure 6.18 shows the expression of genes involved in SL biosynthesis for *Gata6*-KO^{mye}/WT pResMϕ. Gene expression data for the SL metabolic enzymes from the microarray are coloured by fold-change, where red represents upregulation, and green down regulation in *Gata6*-KO^{mye} pResMϕ. These reveal significant changes in gene expression that could potentially lead to the observed lipid dysregulation, and this is described in more detail below.

a) CerS expression directly impacts fatty acyl chain-length in SL in pResMϕ

The six (DH)Cer synthases (CerS), represented in Figure 6.18 transfer variable length fatty acid chains to the free amine group of Sa to form dihydroCer (DH-Cer)[135]. Each of these enzymes utilizes fatty acyl-CoA substrates of various chain lengths resulting in the formation of different DH-Cer subspecies[128,130] (Chapter 1). *CerS2* is responsible for synthesising Cer with chain-lengths C22-26, whilst *CerS4* catalyses formation of C18-20 Cer[135,230]. Thus, an up-regulation in *CerS2* and down-regulation in *CerS4* would in theory lead to an increase of C22-26 Cer. Indeed, this was observed in the microarray. Furthermore, Cer is the precursor for complex SL and changes in its composition are likely to affect DH-SM/SM and DH-GlcCer/GlcCer, *i.e.* the chain length of Cer has a direct impact in the chain length of complex SL, such as SM and GlcCer, thus if there is an increase in CerS of a particular chain length that will be reflected in metabolites downstream of Cer. Accordingly, out of the 6 classes analysed (sphingoid bases, Cer, SM, GlcCer, GalCer, LacCer), C20 SL only increase significantly in DH-GlcCer.

b) Downregulation of Smpd1 in Gata6-KO^{mye} pResMϕ leads to accumulation of DH-SM and SM.

Acid sphingomyelinase (SMPD1) catalyses the breakdown of DH-SM/SM into DH-Cer/Cer, whilst sphingomyelinase synthase (SMS2) mediates the production of SM from Cer in the plasma membrane[231](Chapter 1). Hence the down-regulation of SMPD1 combined with the up-regulation of SMS2, will lead to accumulation of DH- SM/SM, as seen in Figure 6.18.

c) Downregulation of Gba2 in Gata6-KO^{mye} pResMϕ leads to accumulation of DH-GlcCer/GlcCer

Non-lysosomal Glucosylceramidase (GBA2) catalyses the conversion of DH-GlcCer/GlcCer into DH-Cer/Cer in the endoplasmic reticulum, whilst GlcCer synthase (UGCG) mediates the production of DH-GlcCer/GlcCer in the Golgi apparatus. Thus, an up-regulation in *Ugcg* and down-regulation in *Gba2* could explain an increase of DH-GlcCer/GlcCer, which is in accordance with the targeted lipidomics data (section 6.2.2).

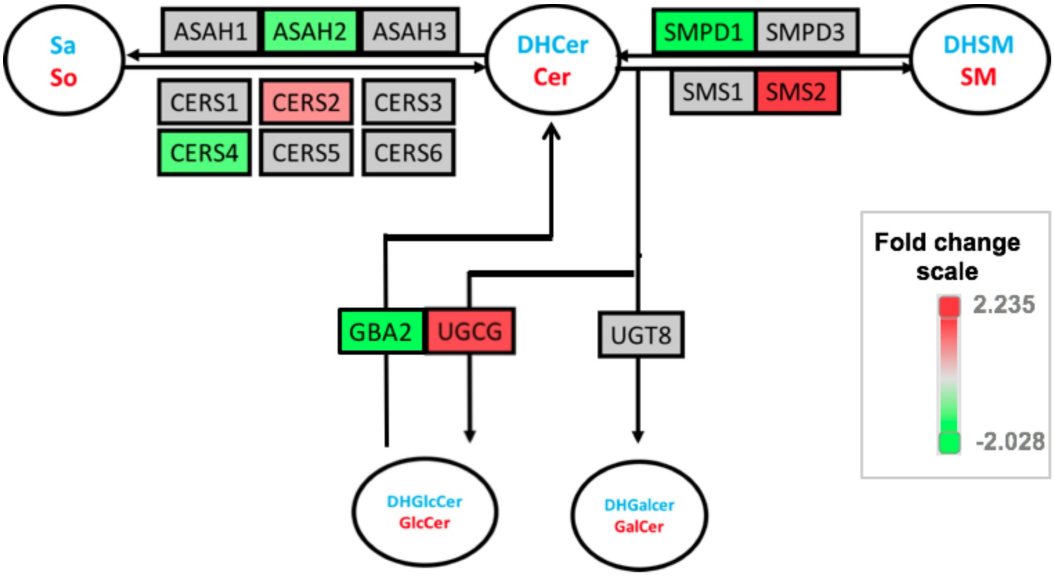


Figure 6.18 Expression of Genes Involved in SL Biosynthesis for the *Gata6*-Ko^{mye}/WT pResMϕ.

Gene expression data for the SL metabolic enzymes from microarray coloured by fold change (WT/KO); Green – Downregulated, Red- Upregulated. Sa is acylated to DH-Cer by (DH)Cer synthases (CerS) and incorporated into more complex DH-SL (Like SM/DH-SM), GlcCer/DH-GlcCer, GalCer/DH-GalCer) or desaturated to Cer followed by headgroup addition. Also included are a number of the catabolic genes, *e.g.*, sphingomyelinases (SMPD), ceramidases (ASAH). The major contributions of each CerS isoform to *de novo* and salvage synthesis of different Cer species are depicted. SL represented in red and DH SL in blue.

6.3 Discussion

In this Chapter I performed a comprehensive analysis of the sphingolipidome of *Gata6*-KO^{mye} and -WT pResMϕ. I used a targeted sphingolipidomic approach to determine the changes of lipid levels upon GATA6 knockdown, and thus explored the role of *Gata6* in regulating Mϕ SL metabolism. I then compared changes in SL gene expression with the lipid changes observed in order to identify the control checkpoints responsible.

SL have central roles in cell structure and function as well as in the pathophysiology of many diseases. Many of the intermediates of SL biosynthesis are highly bioactive and often have antagonistic activities. For instance, Cer promotes apoptosis while S1P can inhibit apoptosis and induce cell growth. Therefore, quantification of SL metabolites, in *Gata6*-KO^{mye} and -WT pResMϕ, is crucial for an understanding the role of GATA6 in SL biology.

The abundance of individual SL species was determined in *Gata6*-KO^{mye} and -WT pResMϕ. This analysis allowed me to compare the relative amounts of the different SL classes. Similar to what had been shown for RAW 264.7, So levels are higher than Sa, and SM is the most abundant, followed by Cer and HexCer[232]. The relative amounts of sphingoid bases vary significantly, and the amount of So is higher than Sa (Figure 6.6C and D, respectively), again agreeing with the literature for RAW264.7[95,156]. Interestingly, in the literature the reported amounts of So are *ca.* 10 times higher than Sa for RAW 264.7 cells, which is 3 times lower than what I found for pResMϕ, Figure 6.6A, whilst for CerP studies have reported that levels of the dihydro forms are considerably lower than those CerP, under basal conditions for Mϕ-like cells[95,156,232]. These differences likely represent cell type specificity. Conversely, in RAW264.7 cells under basal conditions, Cer1P of C16:0, C24:0 and C24:1 are detected, whilst for DH-Cer1P only the C16:0 is found[156]. According to the literature, levels of Cer are 200 times higher than those of DH-Cer in Mϕ-like cells, for the same chain lengths, under basal conditions[232]. Thus, it is possible that the levels of DH-Cer were under the LOD. The majority of these SL are enriched in the plasma membrane[233], however, subcellular organelle lipidomics in Mϕ has shown that headgroup and acyl chain

composition dictate where lipids accumulate[232]. Additionally, their localisation and relative amounts are known to shift upon cell activation. Thus, I believe that a subcellular organelle sphingolipidomics analysis would give an important insight into lipid function in GATA6 deficient pResMϕ.

Herein I have shown that GATA6 deficiency leads to increases in SL levels, where the majority of significantly-altered lipids contain long and saturated fatty acids. This finding supports those from Chapter 5, where I found that *Gata6*-KO^{mye} showed increased levels of SL in pResMϕ, in particular in longer chain HexCer and SM.

The common denominator between the three SM species that changed the most is the fact that other than the plasma membrane they are all mainly enriched in the endoplasmic reticulum, whilst the last is mainly enriched in the mitochondria. Research has shown that membranes rich in these SL are thicker and less fluid, due to the tight packing of their hydrophobic tails and stronger lipid-lipid interactions[234–236].

In regard to fatty acid chain-length, the dysregulation of Cer synthase expression (upregulation of *CerS2* and downregulation of *CerS4*) can potentially be contributing to the differences observed in the significantly increased SL levels. However, further studies are necessary in order to access the contribution of Cer synthases in GATA6 deficient pResMϕ (described in Chapter 8).

It should also be noted that, although acyl chain composition may be influenced by fatty acid availability, the lipid composition can also be a function of additional organelle-specific factors as the fatty acyl profiles. Thus, a targeted analysis of fatty acyls would undoubtedly help explain why the SL with long and saturated fatty acids tend to change more significantly.

Lipidomic analysis of *Gata6*-KO^{mye} and -WT pResMϕ suggests a relationship between the accumulation of different SL classes and the genes responsible for the metabolism of these lipids. As a result, the next Chapter will focus on the knockdown of two selected genes encoding for the enzymes responsible for the degradation of the two SL classes that changed the most significantly, DH-SM/SM

and DH-GlcCer/GlcCer (*Smpd1* and *Gba2)* and do a targeted sphingolipidomic analysis to investigate if the SL accumulation observed in *Gata6*-KO^{mye} pResMϕ can be replicated by downregulating *Smpd1* and *Gba2.* Similarly, I chose a third gene whose expression was decreased by 18 fold in *Gata6*-KO^{mye} pResMϕ, since its downregulation is known to impact lysosomal metabolism (cathepsin E), and can lead to lipid accumulation in pResMϕ[237].

Chapter 7 Downregulation of *Smpd1*, *Gba2* and *Ctse* leads to sphingolipid accumulation

7.1 Introduction

In the previous Chapters I characterised transcriptomic (Chapter 3) and lipidomic (Chapter 5 and 6) changes in GATA6 deficient pResMϕ, which suggested a contribution of SMPD1 and GBA2 in the regulation of SL metabolism in pResMϕ.

GSL, in particular, are ubiquitous membrane components of eukaryotic cells. They participate in various cell recognition events and can regulate enzymes and receptors within the plasma membrane[238]. SMPD1 catalyses the hydrolysis of SM to Cer and choline[239], whilst GBA2 is responsible for the conversion of GlcCer to free glucose and Cer[240]. The downregulation of genes encoding for SMPD1 and GBA2 leads to an accumulation of GSL in eukaryotic cells. GSL accumulation leads to disease in particular sphingolipidoses, where lipid accumulates in cells as a result of specific enzyme deficiencies[241]. For example, downregulation of *Smpd1* results in Niemann-Pick disease, which is characterised by accumulation of SM within the endo-lysosomal compartment[242]. On the other hand, enzymatic dysfunction of glucosidases results in Gaucher's disease[240]. In this disorder, downregulation of *Gba2* in mice results in build-up of GlcCer outside the lysosomes in various tissues (*e.g.* testis and liver) and impairs sperm development and liver regeneration[243].

I also showed in Chapter 3 that the *Ctse* gene was downregulated in GATA6 deficient pResMϕ by 18-fold, when compared to WT controls. The protein this encodes, CTSE, is an intracellular aspartic protease of the pepsin superfamily and is mainly expressed in cells of the immune system, but its location varies depending on the cell type[244]. For example, in mouse Mϕ CTSE is mostly found in endosomal compartments[245]. The tissue distribution of CTSE is controlled by regulation of the *Ctse* gene, so that it is expressed only in certain types of cells. Transcription of *Ctse* depends on the balance between specific transcription factor activities, including as GATA1 and PU1[246]. Lack of *Ctse* expression leads to a lysosomal storage disorder characterised by accumulation of the two major lysosomal membrane sialoglycoproteins LAMP-1 and LAMP-2 and consequently the elevation of lysosomal pH[169]. However, nothing is known regarding the lipid

accumulation originating from this lysosomal disorder. Studies in ARPE-19 cells, show that *Lamp-2* accumulation induces vacuole formation, cell death, cytosolic lipid build-up and decreased exogenous dextran uptake[247]. In particular, when the lysosome's ability to degrade intra- and extracellular material is lost, undegraded material is oxidized, turning to lipofuscin, which is formed by polymeric substances, primarily composed of cross-linked lipid and protein residues due to iron-catalysed oxidative processes[247].

In this Chapter, I will describe the knockdown of *Smpd1*, *Gba2* and *Ctse* in RAW 264.7 cells using a lentiviral shRNA vector and perform a SL analysis on these cells. This is a crucial step in answering the question of whether the lipid changes observed in my experiments is at least in part controlled through altered expression of these genes by GATA6, and it will allow me to compare the results obtained from GATA6 deficient pResMϕ with the targeted knockdowns.

I will begin by producing the lentiviral vectors to knockdown mRNAs of interest, then evaluate how effective they are at down-regulating gene expression. Once the down-regulation is satisfactory, I will engineer three cell lines, with selective knockdown of *Smpd1*, *Gba2* and *Ctse.* I will then quantify the SL levels in these cells, using the methods described in Chapter 7. Additionally, I will assess how the lipid changes observed when specific genes are knockdown compare to lipid changes already seen on *Gata6*-KO^{mye} pResMϕ (Chapter 6).

7.1.1 Aims

- To produce lentiviral vectors to knockdown genes of interest, *Smpd1*, *Gba2* and *Ctse,* in RAW 264.7 cells.
- To determine how the expression of SL metabolism related genes contribute to the pResMϕ lipidome.

7.2 Results

In order to answer the question of whether the lipid changes observed in GATA6 deficient pResMϕ can be explained by the modulation of the expression of *Ctse*, *Gba2* and *Smpd1*, I used lentiviral vectors to generate cell lines where these genes were down-regulated. I then performed a quantitative lipidomic analysis of SL in these cell lines using LC-MS/MS, where the results were compared against a control cell line (RAW 264.7 cells transfected with a non-silencing shRNA).

7.2.1 Validation and characterisation of in vitro knockdown

In order to select the best cell population in which to perform the *in vitro* knockdown of the selected genes, I first analysed expression levels of each gene in different cell populations by qPCR: RAW 264.7 cells, monocyte precursor cells (MOPs) and pResMϕ (as a reference for levels of expression found *in vivo*). The data, presented in Figure 7.1, showed that MOPs do not express *Ctse* . However, the RAW 264.7 cell line expressed all three genes, therefore it was the best candidate for in vitro knockdown of *Ctse* , *Smpd1* and *Gba2*.

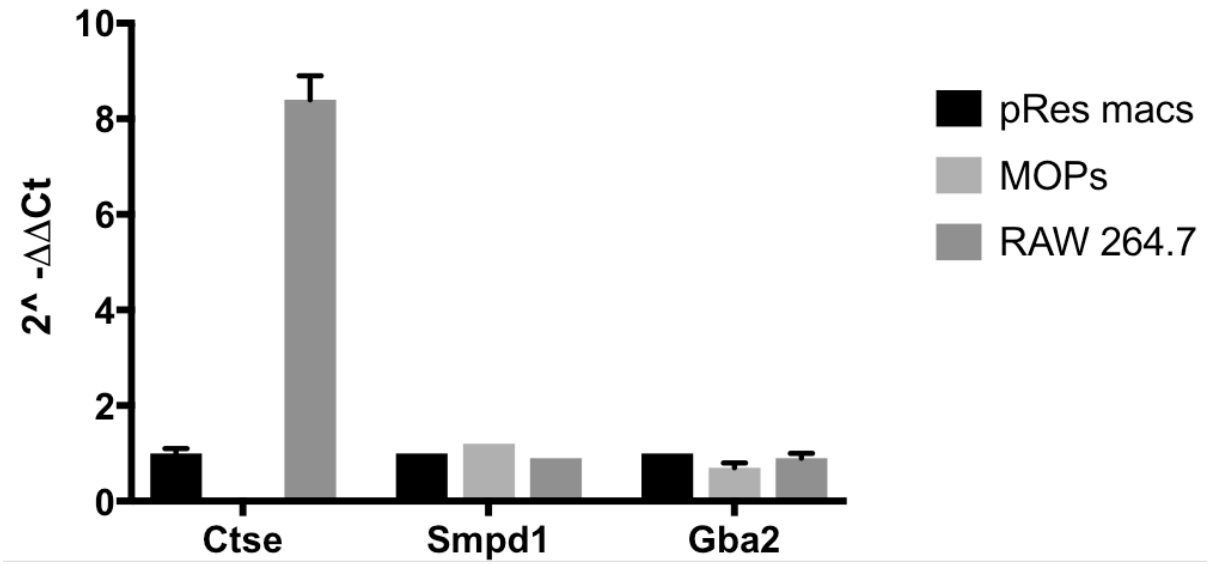


Figure 7.1 *Ctse* , *Smpd1* and *Gba2* in Mϕ.

Ctse, *Smpd1* and *Gba2* mRNA expression compared by qPCR between pRes Mϕ, Monocyte precursors (MOPs) and RAW 264.7 cells. Data shows relative expression, mean ±SEM from one of two independent experiments and is normalized to reflect the number of pResMϕ.

7.2.2 Lentiviral vector induced in vitro knockdown

In order to determine whether these proteins (CTSE, SMPD1 and GBA2) were at least partly responsible for the altered lipidomic phenotype of cells lacking GATA6, I first needed to downregulate the genes responsible for their enconding. The strategy used was RNA interference (RNAi), which is a form of posttranscriptional gene silencing mediated by short double-stranded RNA, known as small interfering RNA (siRNA). These siRNAs can been delivered through lentiviral vectors, to achieve sustained knockdown of target gene expression[248]. Herein, the silencing method used was the production of dsRNA, through sh (short hairpin) RNA, where complementary RNA folds into a hairpin structure and base-pairs to form ds-shRNA[12].

Constructs were designed to silence the expression of *Smpd1*, *Gba2* and *Ctse*, following the strategy described in Chapter 2, section 2.5.2. Three constructs for each gene: non-silencing RNA, *Ctse*, *Smpd1* and *Gba2* (Table 2.4). Figure 7.2 depicts the SU6EW plasmid containing an shRNA insert, used for all constructs. Recombinant plasmids were digested with XhoI and PmeI, and the fragments were identified on 1.5% agarose gel (Figure 7.3). The results of DNA sequencing provided further confirmation of the presence of the recombinant plasmids, indicating that all the shRNA expression plasmids carried the correct sequence.

7.2.3 Efficiency of lentiviral infection in Jurkat cells

Flow cytometry was used to quantify the efficiency of eGFP transfection in Jurkat cells infected with the 12 constructs generated from the SUE6W plasmid, as described in the materials and methods section 2.5.2. Results showed a proportional increase of percentage of infected cells with the volume of lentivirus added to the Jurkat cells (Figure 7.4). All 12 constructs were thus used for transfection of RAW 264.7 cells.

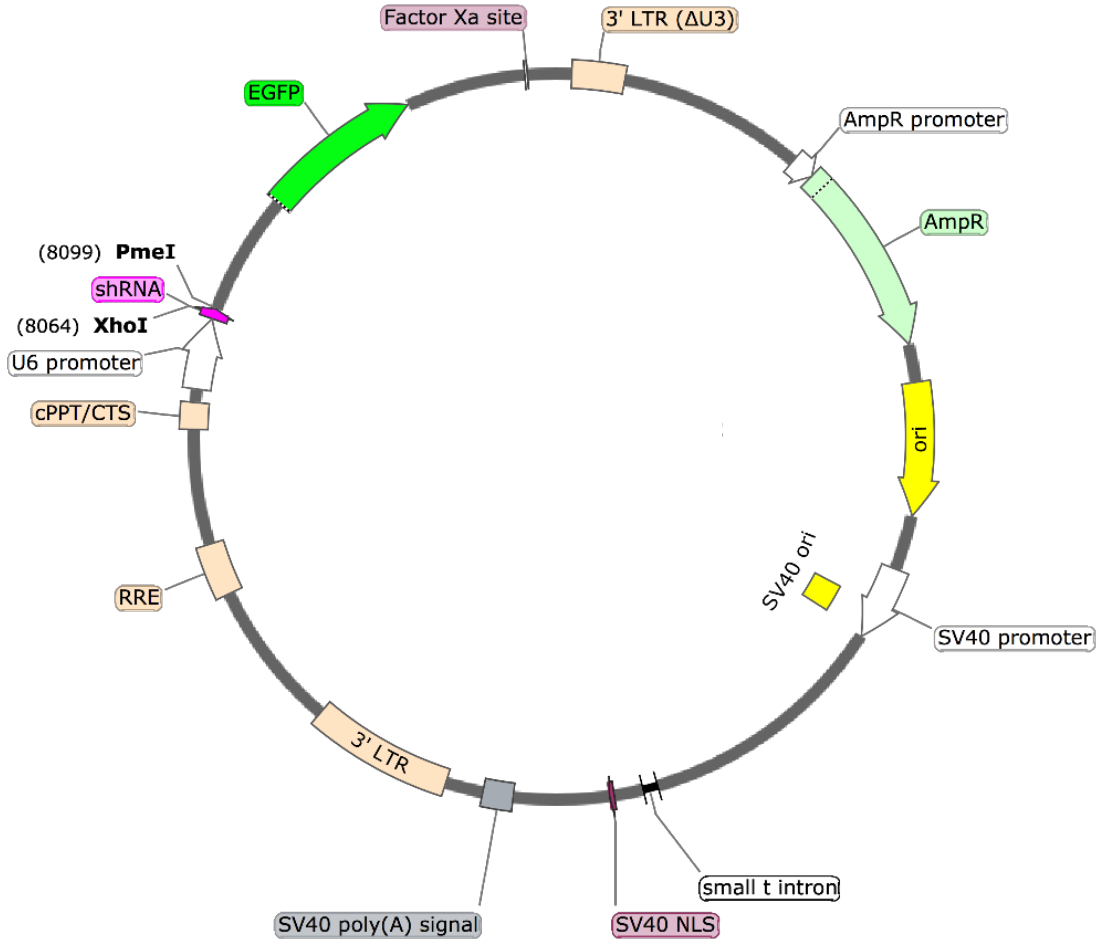


Figure 7.2 Key rgions within SU6EW containing an shRNA insert.

U6 promoter, drives RNA Polymerase III transcription for generation of shRNA transcripts; shRNA, short hairpin RNA; eGFP, green fluorescence protein; Factor Xa site, where factor Xa cleaves after the arginine residue in its preferred cleavage site; AmpR, ampicillin resistence; RRE, Rev response element, allows for Rev-dependent mRNA export from the nucleos to the cytoplasm; cPPT, Central polypurine tract, improves transduction efficiency by facilitating nuclear import of the vector's preintegration complex in the transduced cells; LTR long terminal repeat.

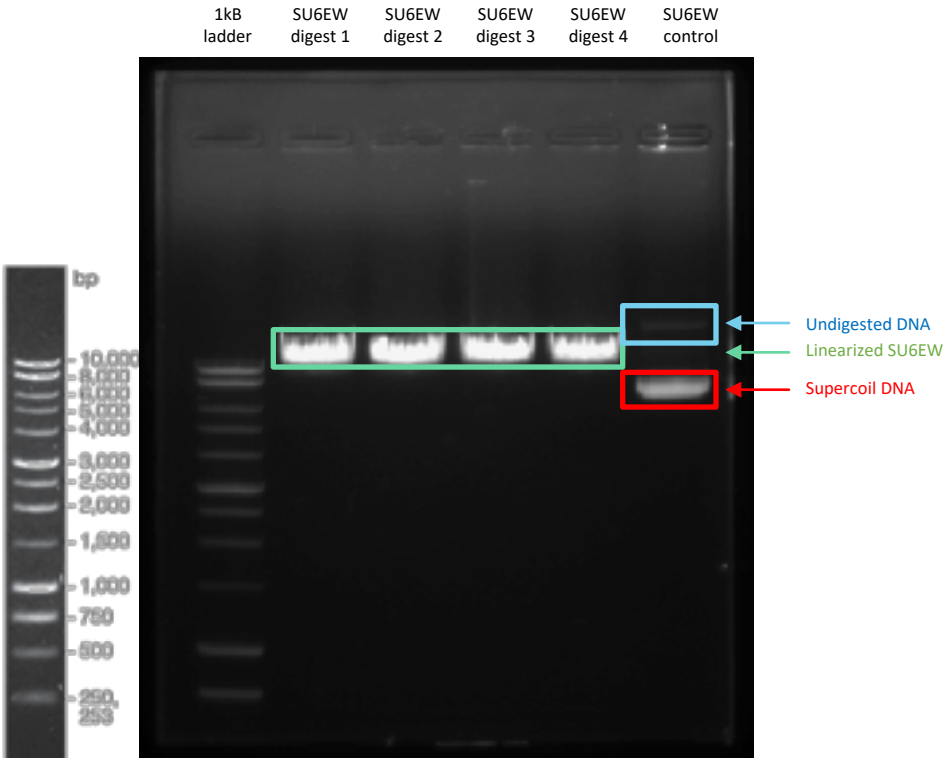


Figure 7.3 Representative agarose gel displaying digestion of plasmids with XhoI and PmeI.

Recombinant plasmids were digested with XhoI and PmeI, and the fragments were identified on 1.5% agarose gel. First lane- 1kB DNA ladder; numbers 2-5 indicate bacterial colonies tested for digestion of plasmids; N – negative control (undigested plasmid).

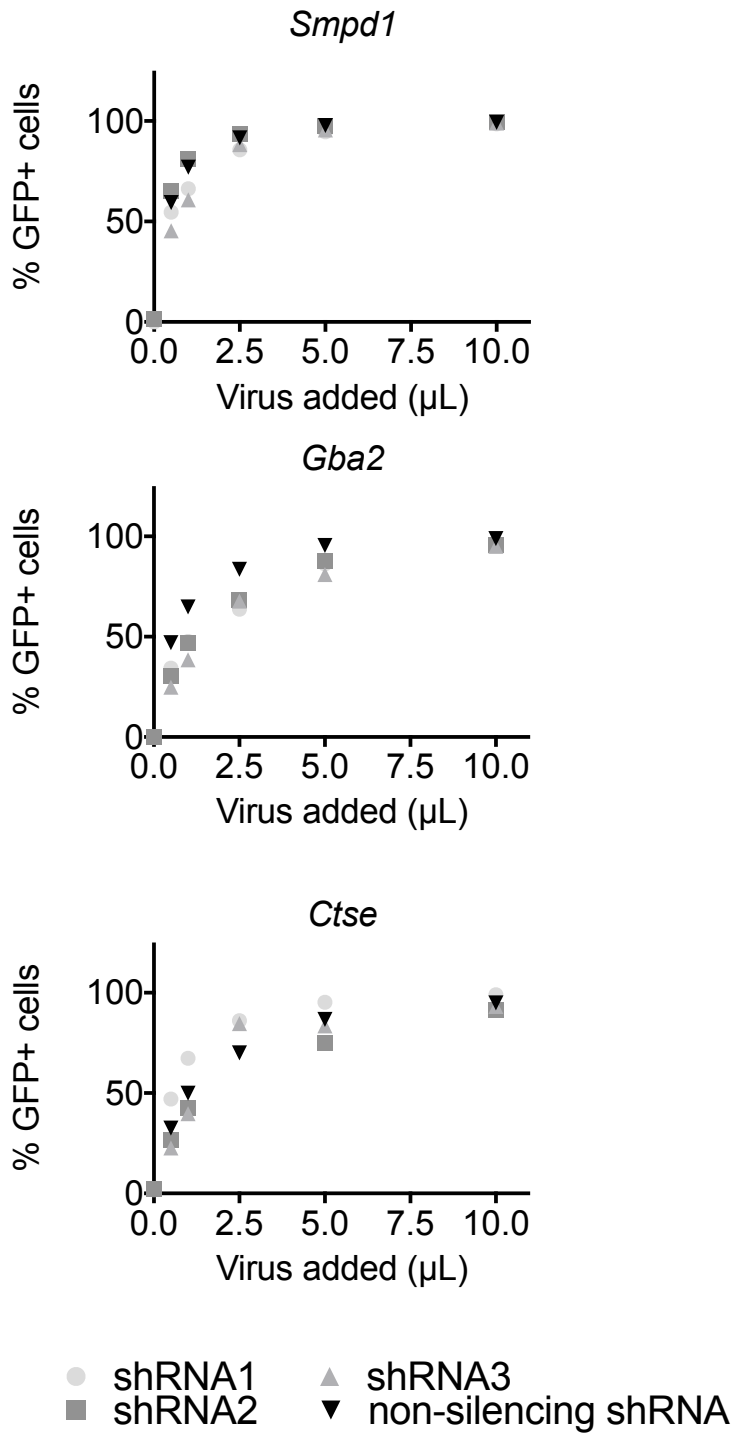


Figure 7.4 Efficiency of eGFP transfection in Jurkat cells increases with viral load.
Scatter plot showing the percentage of Jurkat infected with an eGFP-expressing lentivirus vs the volume (µl) of virus used, 5 days post infection.

7.2.4 Specific lentiviral infection of RAW 264.7 cells

a) shRNA knockdown efficiency

After confirming that all 12 lentiviruses engineered were able to transduce Jurkat cells (Figure 7.4), RAW 264.7 cells were infected *in vitro*. All cell cultures exhibited eGFP$^+$ cells 5 days after transduction (Figure 7.5). The eGFP$^+$ cells were then isolated using Flow sorting and kept in culture for another 5 days, to increase cell numbers, at which point the RNA was extracted and the levels of expression for the genes of interest were analysed (Figure 7.6).

The constructs that maximally reduced the expression of *Ctse*, *Smpd1* and *Gba2* were then selected as follows:

Of the 3 *Ctse* shRNAs, 2 of them showed reduced *Ctse* expression (in RAW 264.7 cells) when compared to the non-silencing shRNA (Figure 7.6A), shRNA1 induced a reduction of 51.6% and shRNA2 induced a 22.15% reduction. However, with regard to the 3 different *Smpd1* shRNAs, only shRNA2 decreased *Smpd1* expression (by 57.9%), Figure 7.6B. For *Gba2*, all 3 shRNAs reduced gene expression, by 36-44.6% (Figure 7.6C).

The gene knockdown levels obtained above were below 60% for all constructs tested, and this was considered to not be sufficient for proceeding to conduct lipidomics. Therefore, I increased the viral load by 2-fold using the 3 constructs that had most strongly reduced expression of *Ctse*, *Smpd1* and *Gba2 (*Figure 7.7*)*

b) Transfection efficiency

The eGFP+ cells were then isolated using flow sorting, then cultured for another 5 days. With the higher amount of virus, there was more than 75% eGFP$^+$ RAW cells at day 5 (Figure 7.8). RNA and lipids were then extracted from these samples to analyse knockdown levels by qPCR and perform a targeted sphingolipidomic analysis.

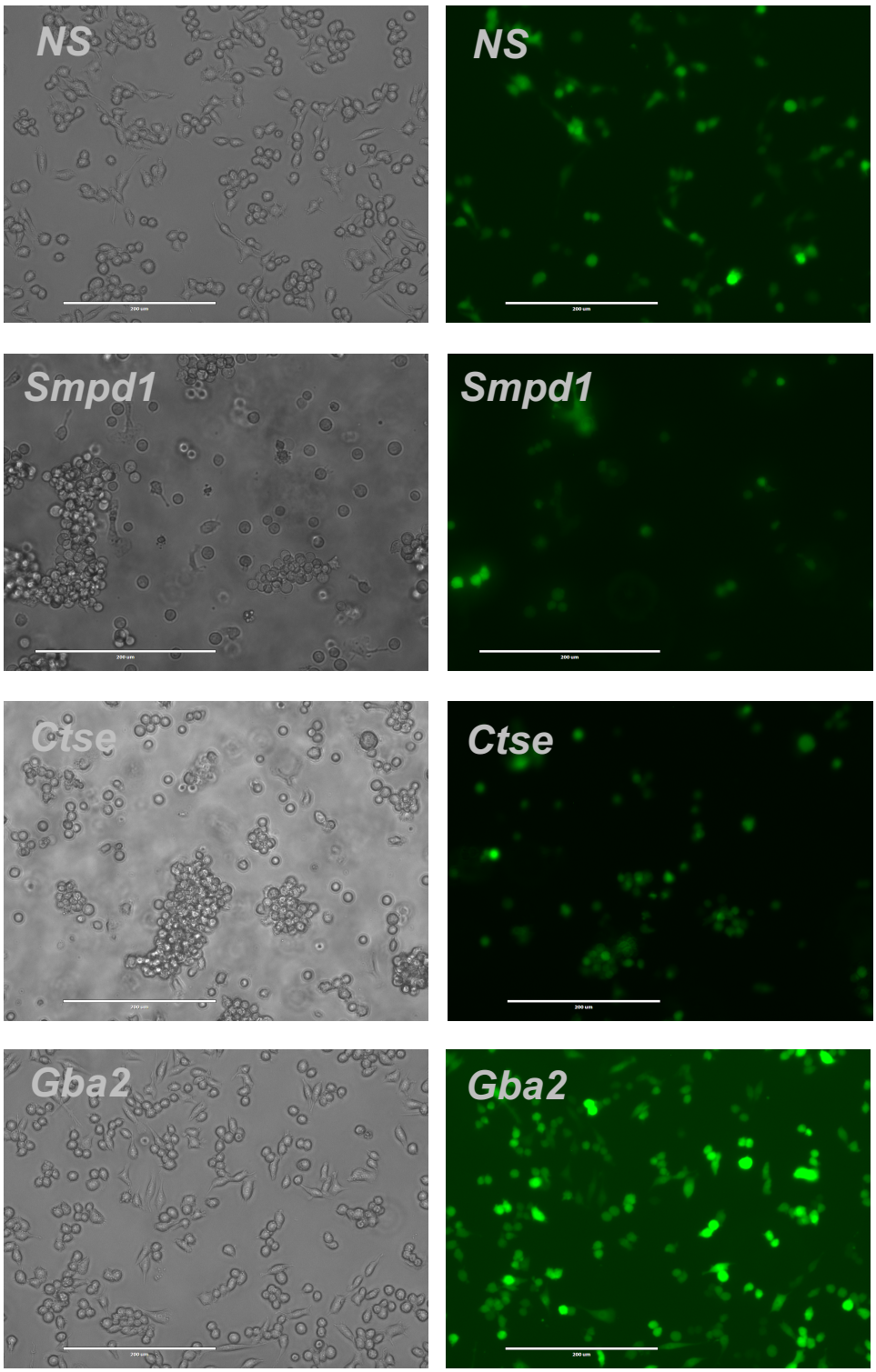


Figure 7.5 eGFP+ expression in RAW 264.7 transfected with 3 different shRNAs and 1 Non-silencing shRNA.

Figure shows representative examples of RAW 264.7 transfected with 3 different shRNAs expressing eGFP 5 days after transduction imaged on an EVOS cell imaging system, excitation 480nm. Left hand side panels, brightfield; right hand side, fluorescence. Bar size = 200 microns

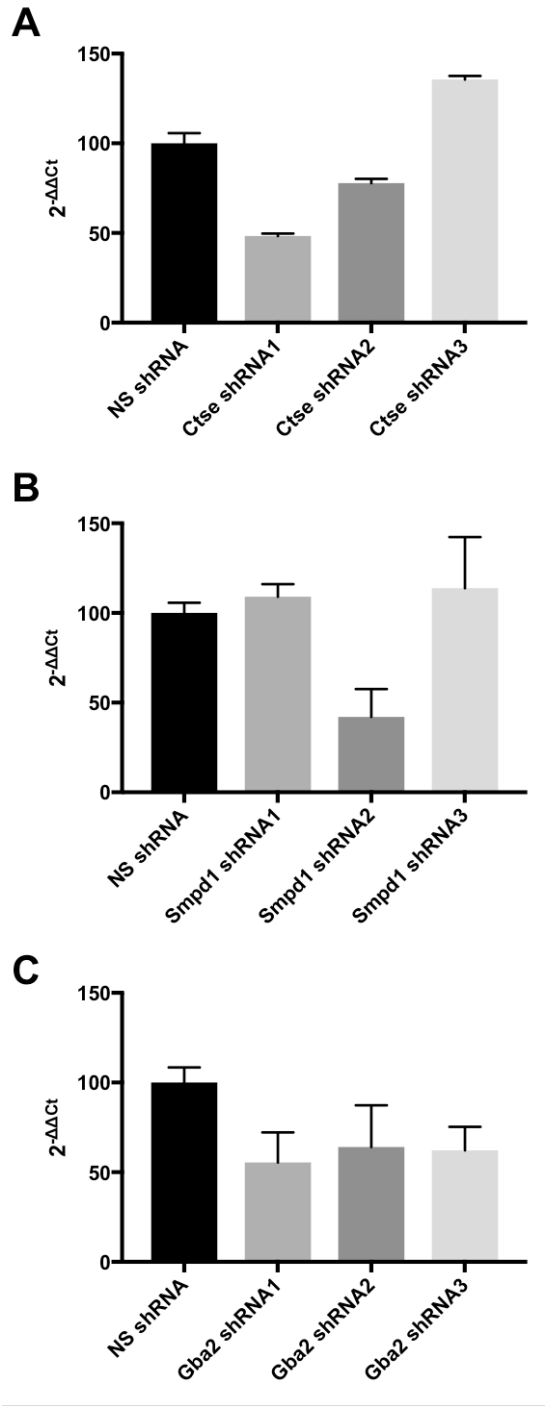


Figure 7.6 Lentiviral knockdown of gene expression in RAW 264.7 cells in vitro.

eGFP+ cells were isolated using Flow sorting 3 days post transfection, and kept in culture for 5 days, RNA was extracted and the levels of expression for the genes of interest were analysed. Gene knockdowns compared by qPCR (Sybr green) in RAW 264.7 transfected with 3 different shRNAs and 1 Non-silencing shRNA. Data shows mean ±SEM from one of two independent experiments. A- *Ctse* , B-*Smpd1*, C- *Gba2*.

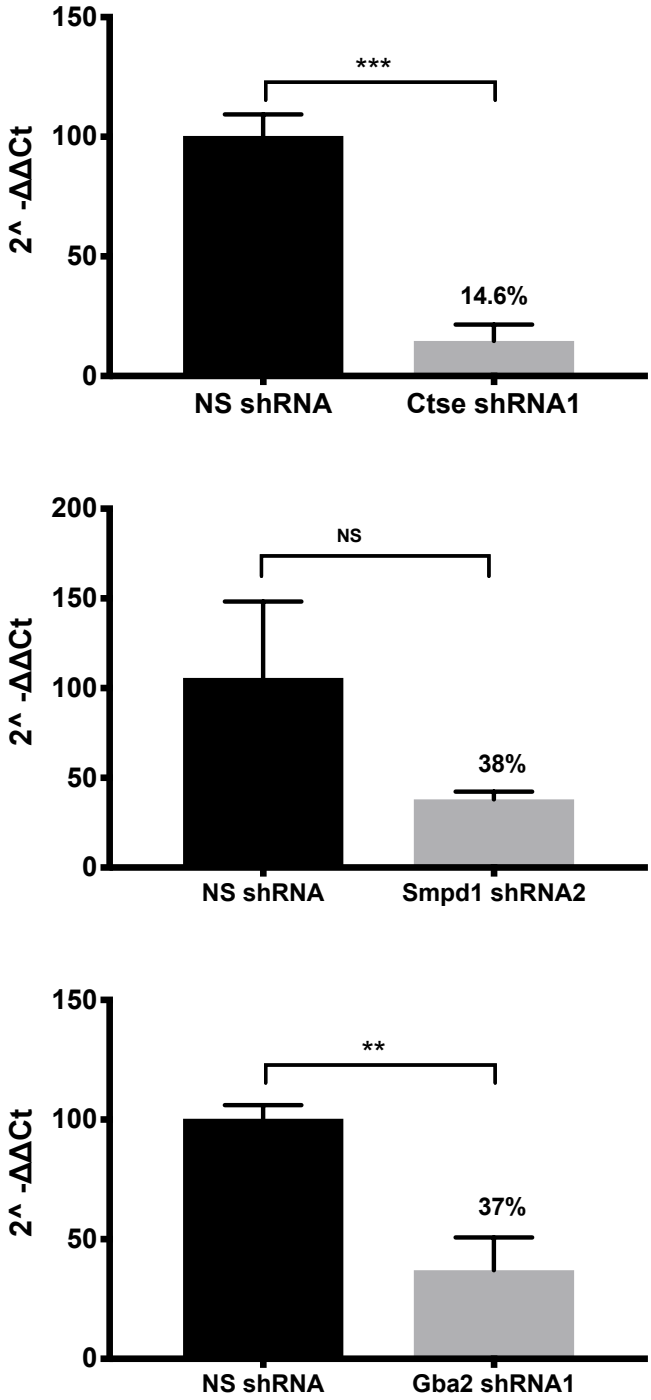


Figure 7.7 Lentiviral knockdown of *Smpd1, Gba2 and Ctse* expression in RAW 264.7 cells in vitro.

eGFP+ cells were isolated using Flow sorting 3 days post transfection, and kept in culture for 5 days, RNA was extracted and the levels of expression for the genes of interest were analysed. Gene expression compared by qPCR in RAW 264.7 transfected with *Smpd1* shRNA and Non-silencing shRNA. Data shows mean ±SEM from one of three independent experiments.

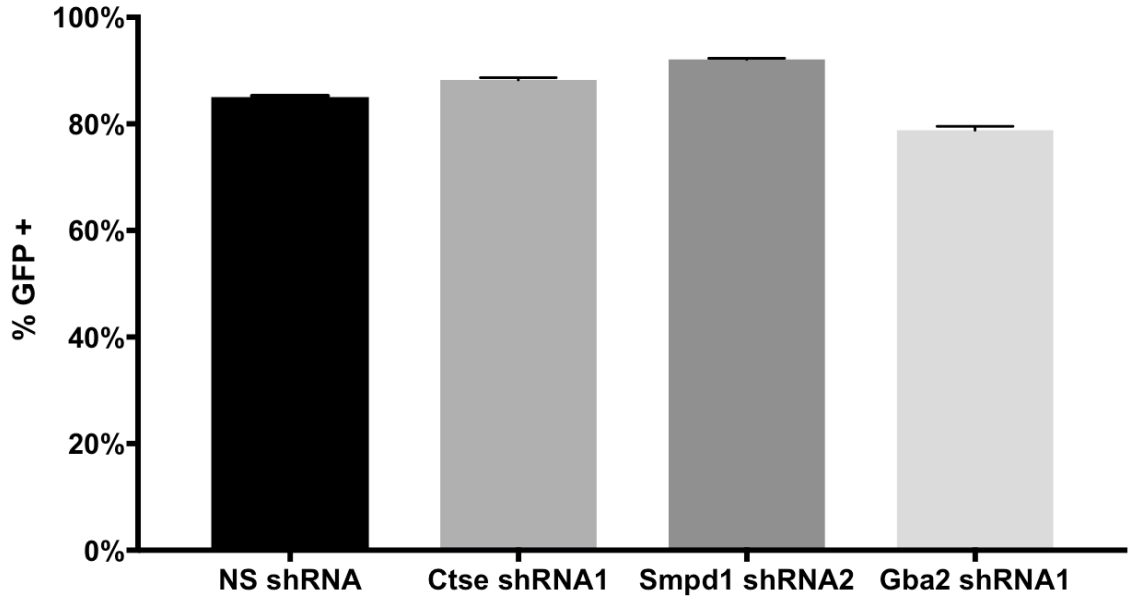


Figure 7.8 Specific lentiviral infection of RAW 264.7 cells in vitro.

eGFP+ cells were isolated using Flow sorting3 days post transfection. Bar plot showing the percentage of GFP positive RAW 264.7 cells, 5 days after infection with an eGFP-expressing lentivirus. Data shown is from the shRNAs with the best knockdown in RAW 264.7 cells (n= 3 per group).

7.2.5 Sphingolipidomic analysis

a) SMPD1 knockdown

Smpd1 expression was knocked down by 62% in RAW 264.7 cells (Figure 7.7). Based on the reactions where the enzyme is involved (namely degradation of SM and DH-SM), this decrease may lead to accumulation of SL, in particular SM (Figure 7.9). However, this will depend both on post transcriptional modifications as well as protein turnover, so caution must be applied here.

Next, lipid extracts from RAW cells transfected with a non-silencing shRNA and *Smpd1* shRNA were analysed in tandem. Complex SL were quantified by LC-MS/MS on a 6500 QTrap mass spectrometer (AB Sciex), using the targeted LC-MS/MS method I developed in Chapter 7 (described in Section 2.6.3).

Using this approach, a total of 86 SL were quantified, of which 13 were found to show significant differences in levels between the two groups (RAW cells transfected with a non-silencing shRNA and a *Smpd1* shRNA), with all of them increasing in SMPD1-deficient RAW264.7. Of the 13 6 were SM, 4 were DH-SM, 2 HexCer and 1 LacCer, with SM and DH-SM changing the most (Figure 7.10**)**.

Structural comparison of the 10 significantly altered SM and DH-SM tended to be saturated or mono-unsaturated (Figure 7.11).

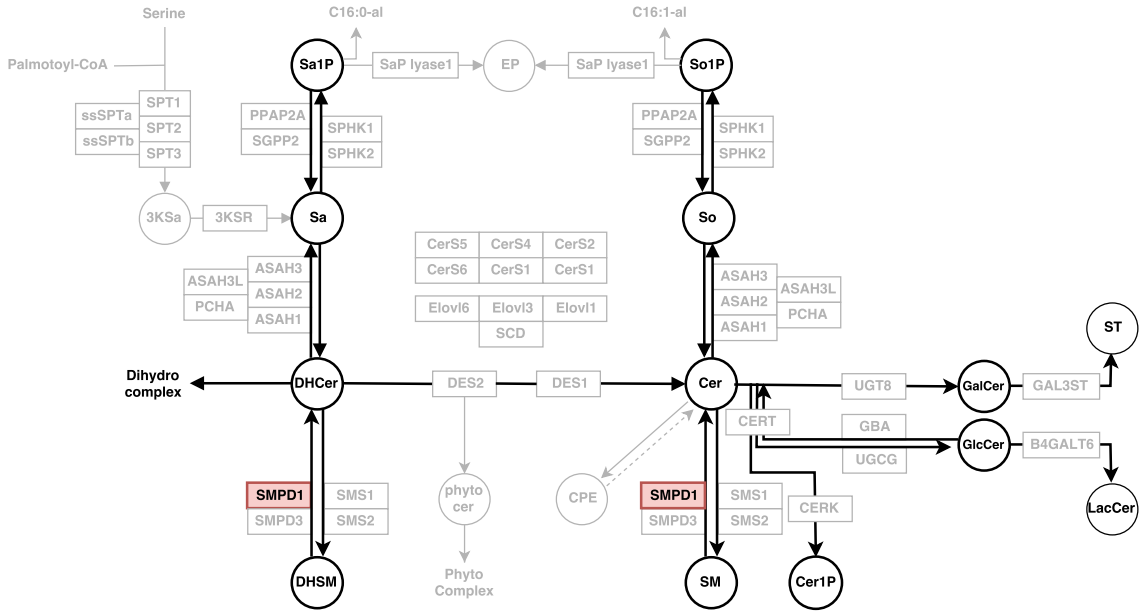


Figure 7.9 Steps of SL biosynthesis.

The figure depicts SL genes (rectangles) and metabolites (circles) that participate in early steps of SL biosynthesis. The metabolites analysed by LC-MS, as previously described, are highlighted in black. The downregulated gene, *Smpd1*, is depicted in red.

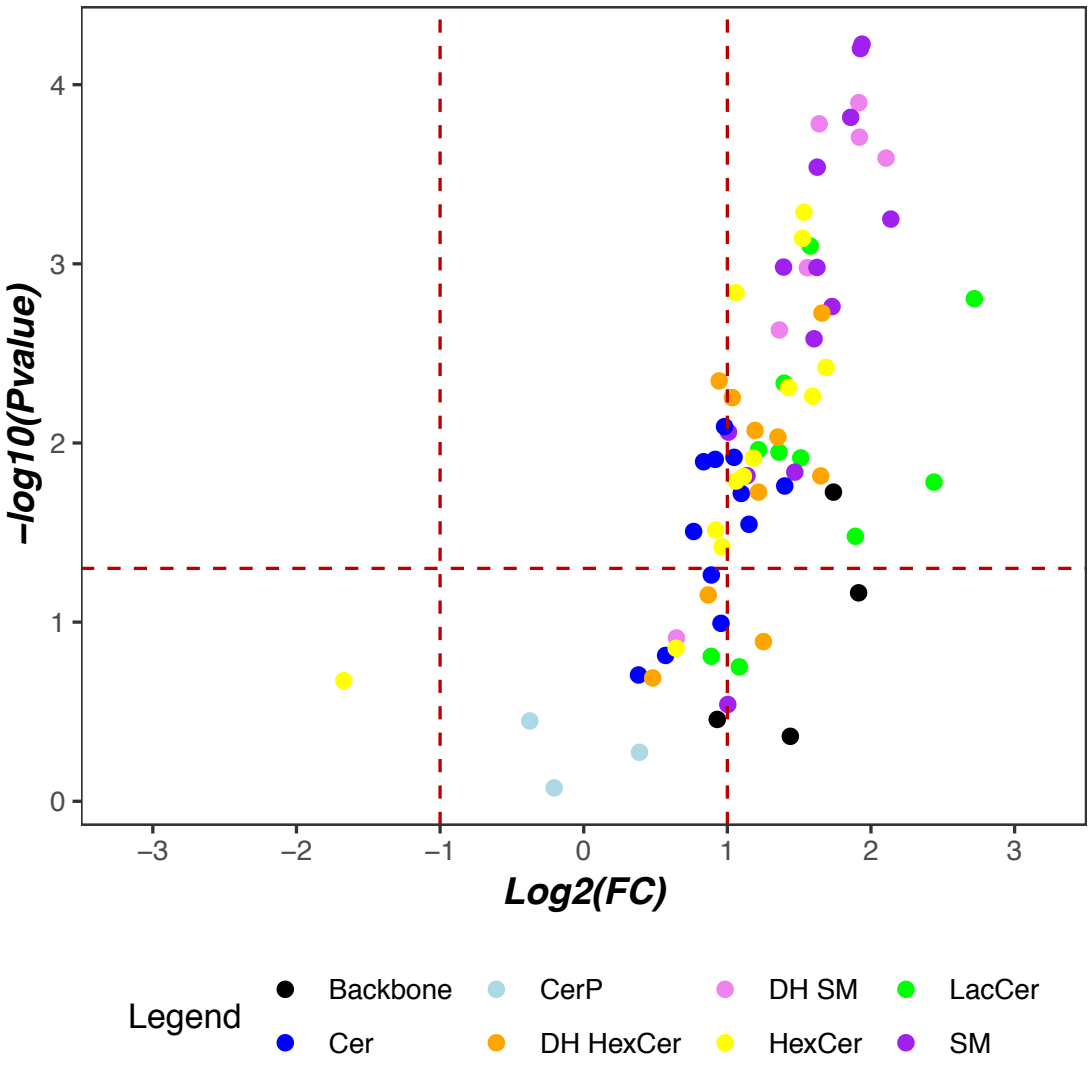


Figure 7.10 SMPD1 deficiency causes an increase in SM in RAW cells.

Lipid extracts from RAW cells transfected with a non-silencing shRNA and a *Smpd1* shRNA. Sa and So backbones (black), Cer (blue), CerP (light blue), HexCer (yellow), DH-HexCer (orange), SM (purple), DH-SM (pink) and LacCer (green) were quantified by LC-MS/MS on a 6500 QTrap mass spectrometer (AB Sciex). Volcano plots representing FC and p-value of lipids in RAW cells transfected with a non-silencing shRNA relative to RAW cells transfected with *Smpd1* shRNA. Statistical significance was determined using the Holm-Sidak method, with alpha= 5.00 %. Each row was analysed individually, without assuming a consistent SD. Differential expression was classified as significant when $P < 0.05$ and $|FC| > 1.5$ (red lines).

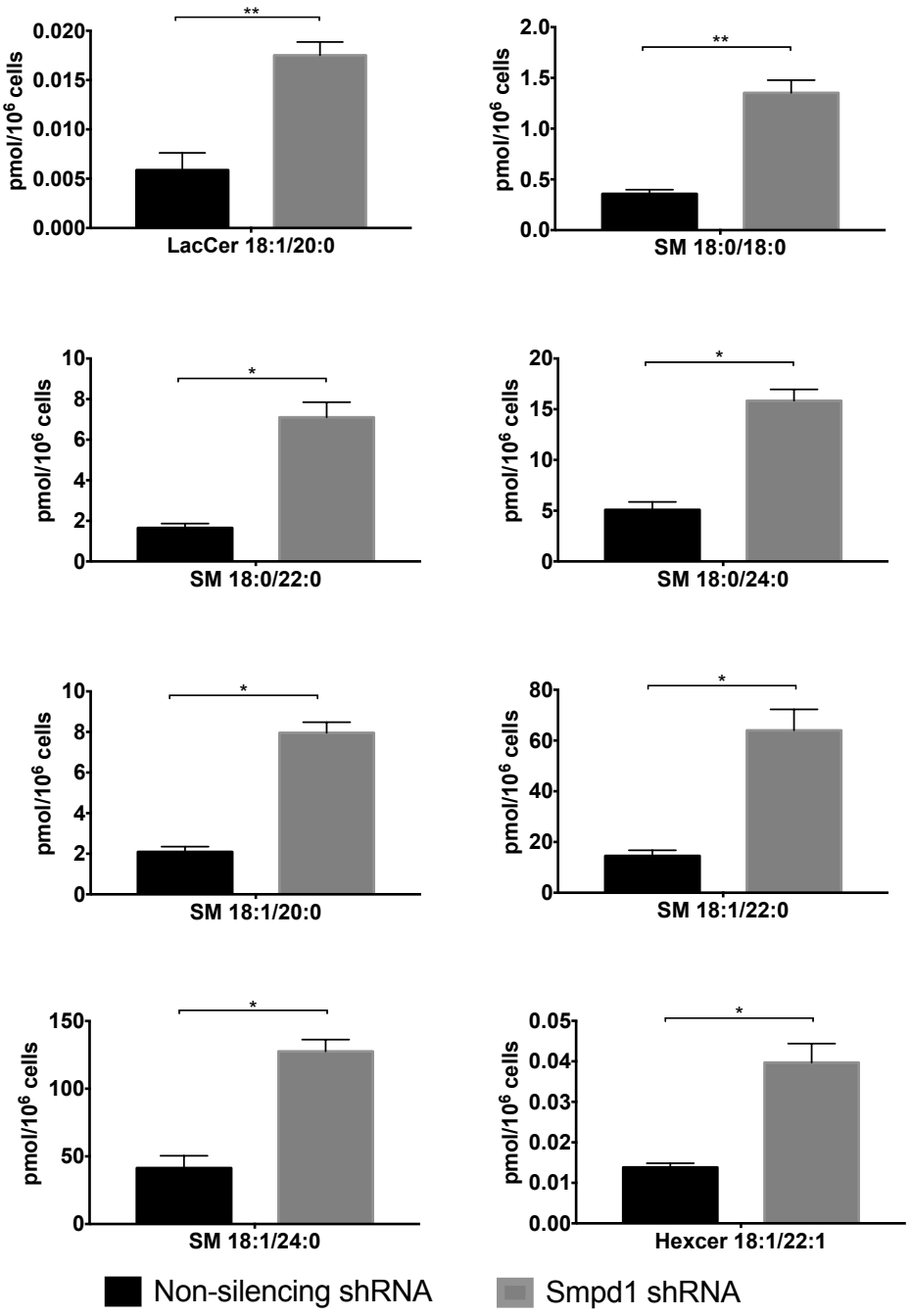


Figure 7.11 Quantitative analysis of significantly altered SL and DH-SL in SMPD1-deficient RAW264.7 cells.

The amounts of these lipids were measured by LC-MS/MS using the internal standard cocktail as described in Chapter 2, section 2.6.2. Lipid extracts from RAW cells transfected with a non-silencing shRNA and a *Smpd1* shRNA were analysed. Complex SL were quantified by LC-MS/MS on a 6500 QTrap mass spectrometer (AB Sciex). Statistical significance determined using the Holm-Sidak method, with alpha=5.000%. Each row was analysed individually, without assuming a consistent SD, the graphs depict the mean and SEM. p-value, *<0.05, **<0.01.

b) Gba2 knockdown

To investigate the role of GBA2 in dysregulation of specific lipid levels, I analysed the sphingolipidome of GBA2-deficient cells. Using the strategy described in Materials and Methods section 2.5.3, I obtained knockdown of 63% expression of GBA2 in RAW 264.7 cells (Figure 7.7). I predicted this would lead to a disruption in SL metabolism, in particular, an accumulation of SL, such as GlcCer (Figure 7.12). To determine this, I conducted a sphingolipidomic analysis of RAW cells transfected with a *Gba2* shRNA and compared these to non-silencing controls. Lipid extracts from RAW cells transfected with a non-silencing shRNA and a *Gba2* shRNA were analysed by LC-MS/MS as described in the previous section. The volcano plot in Figure 7.13, representing FC and p-value of the analysed SL shows that, of the 24 HexCer and DH-HexCer analysed, 14 were significantly increased, of which 7 were HexCer and 7 DH-HexCer.

Comparison of the 14 significantly altered SL and DH-SL in GBA2-deficient RAW264.7 cells (Figure 7.14) revealed no obvious structural similarities in these lipids, other than the fact that they all belonged to the GlcCer class.

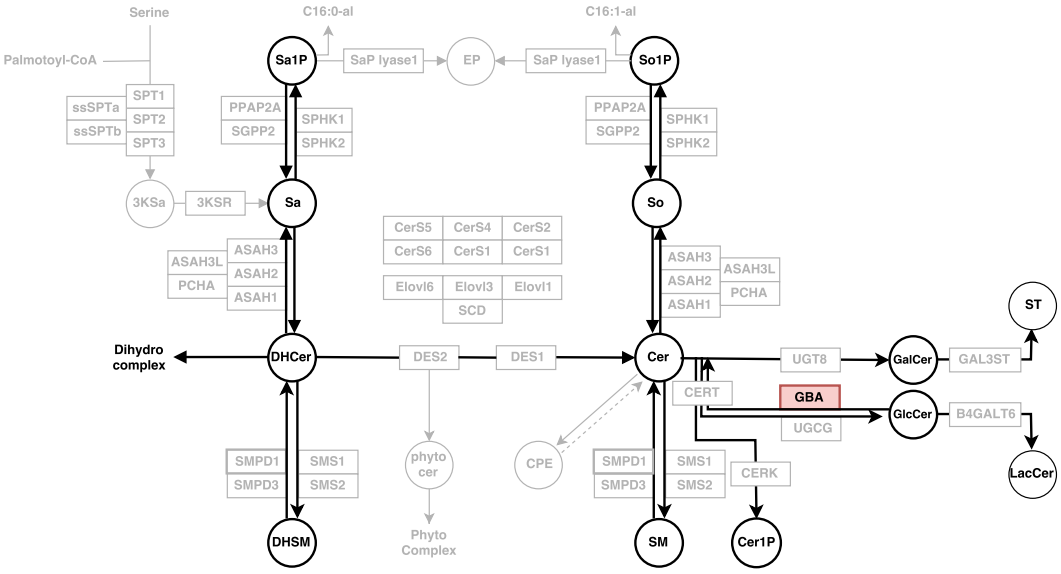


Figure 7.12 Steps of SL biosynthesis.

The figure depicts SL genes (rectangles) and metabolites (circles) that participate in early steps of SL biosynthesis. The metabolites analysed by LC-MS, as previously described, are highlighted in black. The downregulated gene, *Gba2*, is depicted in red.

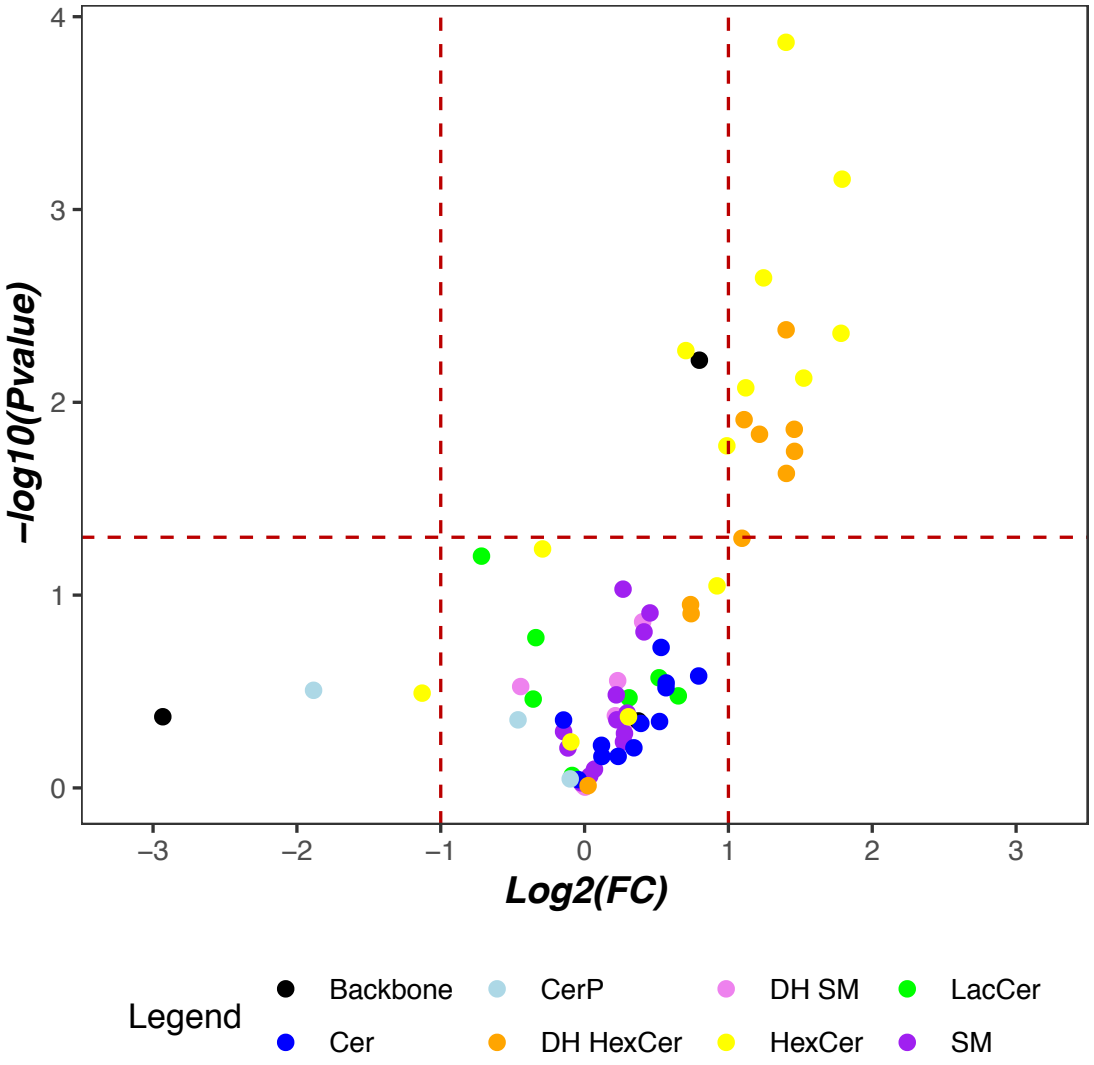


Figure 7.13 GBA2 deficiency causes an increase in GlcCer in RAW cells.

Lipid extracts from RAW cells transfected with a non-silencing shRNA and a *Gba2* shRNA. Sa and So backbones (black), Cer (blue), CerP (light blue), HexCer (yellow), DH-HexCer (orange), SM (purple), DH-SM (pink) and LacCer (green) were quantified by LC-MS/MS on a 6500 QTrap mass spectrometer (AB Sciex). Volcano plots representing FC and p-value of lipids in RAW cells transfected with a non-silencing shRNA relative to RAW cells transfected with *Gba2* shRNA. Statistical significance was determined using the Holm-Sidak method, with alpha=5.00 %. Each row was analysed individually, without assuming a consistent SD. Differential expression was classified as significant when $P < 0.05$ and $|FC| > 1.5$ (red lines).

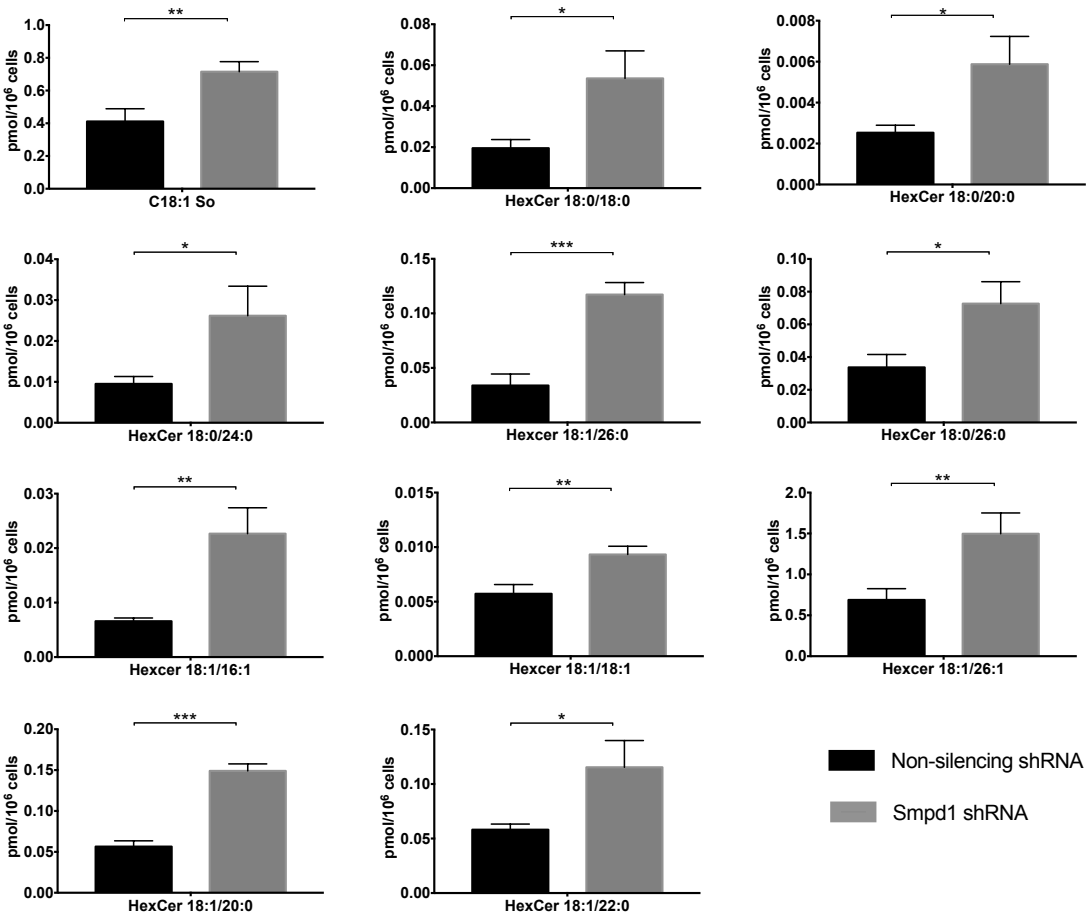


Figure 7.14 Quantitative analysis of significantly altered SL and DH-SL in GBA2-deficient RAW264.7 cells.

The amounts of these lipids were measured by LC-MS/MS using the internal standard cocktail as described in section Chapter 2, section 2.6.2. Lipid extracts from RAW cells transfected with a non-silencing shRNA and a *Gba2* shRNA were analysed. Complex SL were quantified by LC-MS/MS on a 6500 QTrap mass spectrometer (AB Sciex). Statistical significance determined using the Holm-Sidak method, with alpha=5.000% on log transformed data. Each row was analysed individually, without assuming a consistent SD, the graphs depict the mean and SEM. p-value, *<0.05, **<0.01, ***<0.001.

c) CTSE knock-down

To determine the effect of the downregulation of CTSE on lipid levels, I conducted a sphingolipidomic analysis of RAW cells transfected with a *Ctse* shRNA and compared these to non-silencing controls. Lipid extracts were analysed, with complex SL quantified by LC-MS/MS, and data was analysed as described in section 8.3.4.c.

Here, using lentiviral-delivered shRNAs in vitro, I obtained an 85% reduction of *Ctse* expression in Raw264.7 cells. Furthermore, I found that acute shRNA-mediated CTSE reduction led to significant elevations in HexCer and DH-HexCer. The volcano plot in Figure 7.15 shows that, of the 24 HexCer and DH-HexCer analysed, 10 were significantly increased, of which 7 were HexCer and 5 DH-HexCer.

Comparing the structures of the 10 significantly-altered SL and DH-SL in CTSE-deficient RAW264.7 cells (Figure 7.16) revealed structural similarities, since all lipids increased were found to be either HexCer or DH-HexCer.

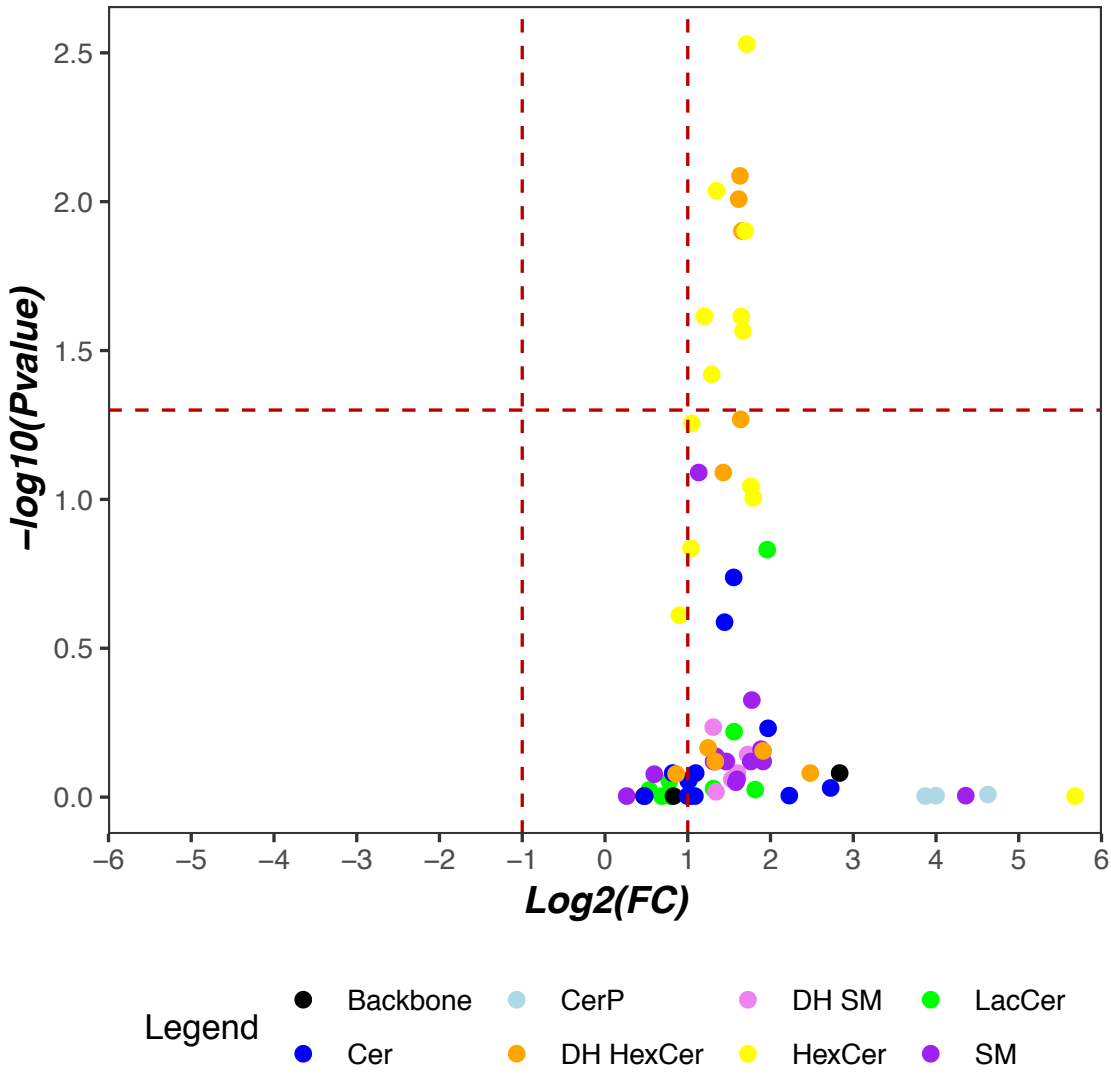


Figure 7.15 CTSE deficiency causes an increase in GlcCer in RAW cells.

Lipid extracts from RAW cells transfected with a non-silencing shRNA and a *Ctse* shRNA. Sa and So backbones (black), Cer (blue), CerP (light blue), HexCer (yellow), DH-HexCer (orange), SM (purple), DH-SM (pink) and LacCer (green) were quantified by LC-MS/MS on a 6500 QTrap mass spectrometer (AB Sciex). Volcano plots representing FC and p-value of lipids in RAW cells transfected with a non-silencing shRNA relative to RAW cells transfected with *Ctse* shRNA. Statistical significance was determined using the Holm-Sidak method, with alpha=5.0 %. Each row was analysed individually, without assuming a consistent SD. Differential expression was classified as significant when $P < 0.05$ and $|FC| > 1.5$ (red lines).

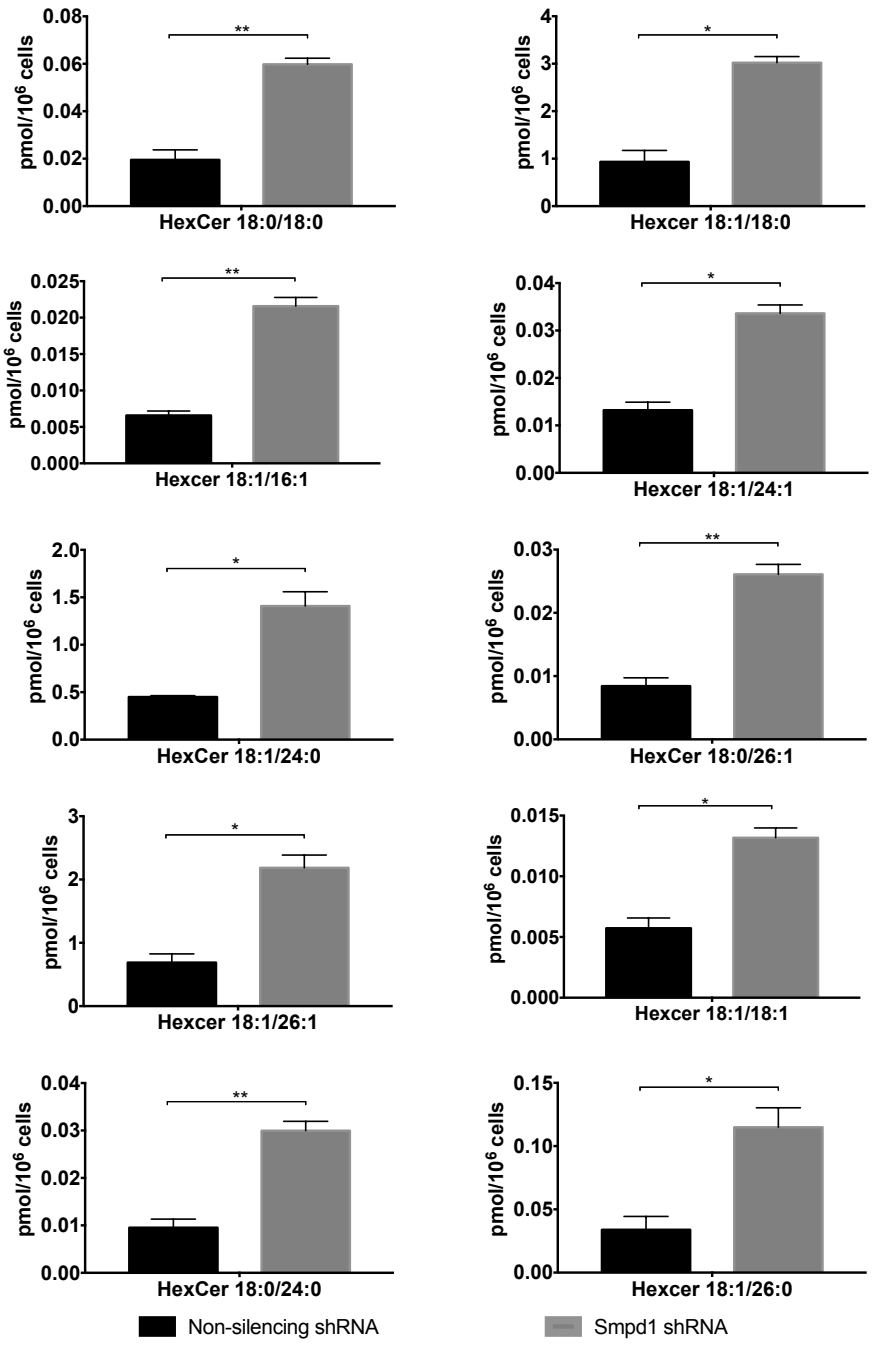


Figure 7.16 Quantitative analysis of significantly altered SL and DH-SL in CTSE-deficient RAW264.7 cells.

The amounts of these lipids were measured by LC-MS/MS using the internal standard cocktail as described Chapter 2, section 2.6.2. Lipid extracts from RAW cells transfected with a non-silencing shRNA and a *Ctse* shRNA were analysed. Complex SL were quantified by LC-MS/MS on a 6500 QTrap mass spectrometer (AB Sciex). Statistical significance determined using the Holm-Sidak method, with alpha=5.000% on log transformed data. Each row was analysed individually, without assuming a consistent SD, the graphs depict the mean and SEM. p-value, *<0.05, **<0.01.

7.3 Discussion

GSL are degraded in a strictly sequential pathway by endolysosomal enzymes with the assistance of small lipid binding proteins, the so-called 'sphingolipid activator proteins'. Inherited metabolic defects in any one of the steps, where these proteins are involved, may lead to a block of degradation and accumulation of the corresponding substrates, resulting in a group of lysosomal storage diseases, termed sphingolipidoses[250].

In this study, I used lentiviral vectors to induce specific knockdown of *Smpd1*, *Gba2* and *Ctse* in RAW 264.7 cells, combined with a targeted sphingolipidomic approach to determine the role of these genes in regulating SL metabolism in Mϕ.

Initially, this experiment was designed to test the knockdown conditions that would later be reproduced *in vivo*. However, due to time constraints it was not possible to proceed to *in vivo* experiments.

7.3.1 *Smpd1* knockdown

SMPD1 occupies an important position in SL catabolism, catalysing the hydrolysis of SM to Cer and phosphorylcholine[242]. Insights into the importance of SMPD1, come from the study of the human lysosomal storage disorder Niemann Pick Disease Types A (NPD A) and B (NPD B)[251]. The accumulation of excessive levels of SM in NPD patients was initially described by Klenk in 1934[252]. Since then, studies using *Smpd1*$^{-/-}$ mice have shown age-dependent accumulation of SM and development of Niemann-Pick syndrome A or B, depending on the residual activity of SMPD1.

Despite the multitude of studies implicating the SMPD1/Cer pathway in important cellular processes, few have addressed how altering gene expression modulates the accumulation of different lipid species from this large family. The quantitative analysis of complex SL in RAW cells transfected with *Smpd1* shRNA relative to RAW cells transfected with a non-silencing shRNA by LC-MS/MS, revealed an increase not only in SM and DH-SM, but also in LacCer and HexCer lipids (Figure 7.11). These data seem to largely agree with the findings previously

reported by M. Koberlin, *et al.* in *Smpd1* knockdowns in RAW cells[98]. However, in this study they observe a significant decrease in Cer levels that my data did not show, which could be due to the fact that the method used by M. Koberlin, *et al.* for lipid quantification was less specific, which could lead to significant differences in results. Nevertheless, this change in Cer levels was expected, as previously mentioned, SMPD1 is responsible for the conversion of SM into Cer, thus less degradation of SM should have an impact on the levels of Cer. According to published data, accumulation of SM, is positively correlated with cholesterol levels in the circulation as well as the retention of cholesterol in the plasma membrane[253,254], and the regulation of cell migration[255], and can potentially be the cause for some of these phenotypical changes when observed in *Gata6*-KO^{mye} pResMϕ[69].

7.3.2 *Gba2* knockdown

Gba2 is a β-glucosidase, responsible for the cleavage of GlcCer (to glucose and Cer). GlcCer is not only a main precursor for other GSL but it is also a relevant intermediate in the pathways of GSL metabolism and plays key roles in the regulation of cell physiology. The molecular mechanisms by which GlcCer regulates cellular processes are unknown, but might involve changes in membrane biophysical properties and formation of lipid domains[256].

Knockout studies have shown that GBA2-deficient mice accumulate GlcCer in tissues that have the greatest expression of GBA2, *i.e*. in the brain, spleen and testis[243,257]. This indicates an active role for GBA2 in metabolising GlcCer in these sites. Accumulation of GlcCer leads to male infertility due to the formation of aberrant sperm (large round heads , abnormal acrosomes, and defective mobility), and delayed liver regeneration after partial hepatectomy[258,259]. The mechanisms responsible for these phenotypes remain unknown because the signalling function of GlcCer is not well understood, in particular, the fate of non-lysosomal GlcCer[243].

My work showed that the knockdown of *Gba2* in RAW 264.7 cells disrupted SL metabolism, leading to accumulation of GlcCer and DH GlcCer (Figure 7.13),

whereas total levels of sphingoid bases, Cer, SM, and LacCer were not significantly affected (Figure 7.13).

This type of SL accumulation (GlcCer and DH-GlcCer), has been shown to lead to the same disruptions in cell metabolism as those observed in *Gata6*-KO^{mye}, such as alternative activation of Mϕ activation, tissue inflammation and proinflammatory cytokine production (IL-6 and TNF-α)[71,251,260]. However, there is no evidence that the same mechanisms are responsible for the phenotypical changes observed in *Gata6*-KO^{mye}.

7.3.3 *Ctse* knockdown

CTSE is an intracellular aspartic protease of the endolysosomal pathway, predominantly expressed in cells of the immune system. It has been implicated in several physiological and pathological processes however, so far the functional importance of CTSE is not completely understood[261]. Studies using knockout mice have provided some information regarding the contribution of CTSE to physiology. Mϕ from CTSE-deficient mice exhibit a form of lysosome storage disorder, exhibiting the accumulation of the two major lysosomal membrane sialoglycoproteins LAMP-1, LAMP-2 and LC3 and the elevation of lysosomal pH[169]. However, nothing is known regarding the lipid dysregulation originating from this lysosomal disorder. Herein, I analysed cells where *Ctse* was downregulated, and observed that the knockdown of *Gba2* and *Ctse* in RAW 264.7 cells led to a disruption in SL metabolism, with similar effects, *i.e.* accumulation of GlcCer and DH GlcCer (Figure 7.15). However, there was no evidence of a preferential accumulation of the same lipid species that were significantly increased in *Gata6*-KO^{mye} cells. In fact, all SL species increased to some degree, but only the changes in GlcCer and DH GlcCer levels were significant. This result was quite unexpected as the only information in the literature regarding CTSE and lipid metabolism referred only to a lysosomal disorder, without any mention of its influence in lipid metabolism. It can be speculated that this accumulation could potentially be due to the dysregulation of lysosome function caused by the accumulation of LAMP-1 and LAMP-2.

Even though CTSE is not a lysosomal cathepsin, as it is only present in endosomes, it still plays an important role in lysosomal function, with its deficiency leading to impaired fusion of autophagosomes with lysosomes and leading to increased oxidative stress[262]. The phenotype of CTSE-deficient Mϕ is similar to that of cells lacking lysosomal proteins. For example, LC3 protein levels are elevated in cells lacking lysosomal proteins, in phenotypes found in neuronal ceroid lipofuscinosis, and Niemann-Pick Type C disease, where there is accumulation of SL [262]. However, I was not able to find literature indicating that GlcCer and DH-GlcCer would be more affected by the downregulation of *Ctse* compared with other lipid species. Thus, these findings are quite novel and this is a topic that requires further investigation.

7.3.4 GlcCer and SM accumulation

Overall, the lipidomics results from CTSE/GBA2 and SMPD1 deficient cells combined, support my hypothesis that GATA6 regulates SL metabolism in peritoneal Mϕ by altering the expression of these genes. This further suggests that these pathways and lipids could be essential regulators of the pResMϕ phenotype in mice.

Future functional experiments designed to address this hypothesis, focussing on restoring the WT phenotype in *GATA6*-KO^{mye} by modulating CTSE, GBA2, and SMPD1 activity, will be discussed in the next Chapter.

Chapter 8 General discussion

ResMϕ are distributed throughout the body and play important roles in pathogen detection and homeostasis[263]. In tissues, these cells exhibit distinct phenotypes and activities, yet the precise mechanisms that regulate these are still largely unknown[65]. pResMϕ perform an essential role in the regulation of infectious and inflammatory diseases. Although recent progress on murine pResMϕ has revealed multiple aspects on their origin and mechanisms involved in their maintenance in this compartment, little is known on their characteristics in homeostasis. In this thesis I investigated whether the myeloid deficiency of GATA6 was associated with alterations in lipid metabolism, at a transcriptional and molecular level in pResMϕ. GATA6 is a master regulator of the phenotype and function of pResMϕ, whose deficiency results in dysregulated proliferative renewal during homeostasis and altered inflammatory responses, associated with delays in resolution[69]. However, whether GATA6 can directly influence lipid metabolism was not known before now, apart from some preliminary findings in a transcriptional dataset that showed several differentially expressed genes involved in lipid metabolism[69]. Among the first genes identified were ABCA, ALOX15, SPTLC2, FABP4 and LDLR. Such changes in gene expression are presumably linked to alterations in lipid metabolism related pathways.

Lipids play multiple essential roles in cell signalling and subcellular structure. Mϕ are required to deal with the lipids they obtain through the phagocytosis of dying cells, however, it is clear from the transcriptional profiles of different ResMϕ that some, namely the alveolar Mϕ, pResMϕ, and Kupffer cells, are more highly specialised in this function than others[57,66,264]. This is a crucial function of these Mϕ as defective or absent lipid metabolism can result in several pathologies including inflammatory diseases and abdominal cancers in the peritoneal cavity. Furthermore, progress of pathologies can also lead to the recruitment of inflammatory monocyte-derived Mϕ which, in turn, take up lipids and subsequently need to be able to process them, potentially having dramatic effects on disease progression and patient outcome such as in atherosclerotic plaques. Thus, it is evident that Mϕ play key roles in lipid metabolism and in the pathogenesis of lipid-related diseases. However, it is also clear that this is not yet fully understood. Thus, there is still much to learn regarding the roles of Mϕ

particularly in lipid metabolism. Importantly, as the understanding of Mϕ increases, both in terms of their common and tissue-specific features, it is easier to design new tools to address such specific questions. Hence a systematic analysis of the total lipid structure of a cell or organism, can reveal novel avenues for therapeutic intervention and diagnosis.

The data presented within this thesis has contributed to the growing knowledge on Mϕ lipid biology. I found that the GATA6 transcription factor regulates the expression of lipid metabolism related genes in pResMϕ, in particular, SL metabolism-related genes (Figure 8.1, transcriptomic analysis), as evidenced by the dysregulation of SL metabolism-related pathways and causal mechanisms (Chapter 3). A wide range of genes with roles in lysosomal function were found to be dysregulated (Chapter 3). I next optimised the LC-MS conditions necessary to perform an untargeted lipidomic analysis of Mϕ-like cells, using pResMϕ and RAW 264.7, and thus I found that the lipidome of these 2 cell types is significantly different, which suggests considerable differences in phenotype (Chapter 4). Specifically, RAW 264.7 cells showed increased levels of lipids that are known to lead to a pro-resolving phenotype (steroid lipids), whilst pResMϕ show the opposite, levels of bile acids are significantly higher, which implies a more pro-inflammatory phenotype[198–200].

After optimising the LC-MS conditions in Chapter 4, I explored the metabolic defect caused by the KO of *Gata6*, by comparing the lipidome of *Gata6*-KO^{mye} and WT pResMϕ, using the high-resolution LC-MS based global lipidomics approach established in Chapter 4. This revealed that putatively identified SM and HexCer lipids were significantly increased in *Gata6*-KO^{mye} cells (Figure 8.1, untargeted analysis) (Chapters 5).

Mapping the expression results onto the known SL metabolic pathway showed that 13 of the 20 genes were dysregulated in GATA6-defficient cells (Figure 8.2), consistent with the fact that that SL and GPL levels were increased. The genes induced by GATA6 were associated with the *de novo* synthesis of Cer (Sptlc2 and CerS2) and their downstream processing into other key SL including SM by *Cola3bp*, and GluCers by *Ugcg*. In contrast, genes involved in the

degradation of SM (*Smpd1* and related genes) and GlcCer (*Gba2*) were predominantly downregulated. This significant and consistent transcriptional pattern across the network suggested increased levels of SM and GlcCer, as indeed observed in the untargeted lipidomics data lipidomics data (Figure 8.1).

To directly quantify SL in pResMϕ, I then developed and validated a quantitative assay which I applied to determine the changes of lipid levels upon *Gata6* KO, and thus explored the role of GATA6 in regulating SL metabolism in Mϕ. Using this approach, I found that HexCer and SM lipids were significantly increased in GATA6 deficient cells, which suggests a relationship between the accumulation of different SL classes and the genes responsible for the metabolism of these lipids (Chapter 6) (Figure 8.1, targeted sphingolipidomic analysis). SM and HexCer are involved in cell proliferation, apoptosis and metabolic regulation[265,266]. Their accumulation in tissues has been previously implicated in a plethora of health complications such as development of neurological dysfunctions, atherosclerosis, diabetes, and heart failure[242]. Therefore, I next examined the ability of GATA6 to modulate their levels, by regulating the expression of genes involved in SL metabolism (Chapter 7). The increase in SL levels linked with downregulation of *Gba2*, *Ctse* and *Smpd1*, observed during the microarray analysis, was confirmed by supressing these pathways, in RAW 264.7 cells, by means of knockdown (shRNA) methodologies. Using this approach, I showed that downregulation of *Smpd1*, leads to an increase in some of the DH-SM and SM species affected by GATA6 deficiency (Figure 8.3). Some difference in the SL species present was expected considering the results obtained from the untargeted lipidomic analysis of pResMϕ and RAW 264.7 cells (Chapter 4), where 66% putatively identified SL were found to be significantly dysregulated, with more species present in RAW 264.7. SMPD1 is part of the SL *de novo* pathway, and it was expected to have a direct impact on SM levels[242].

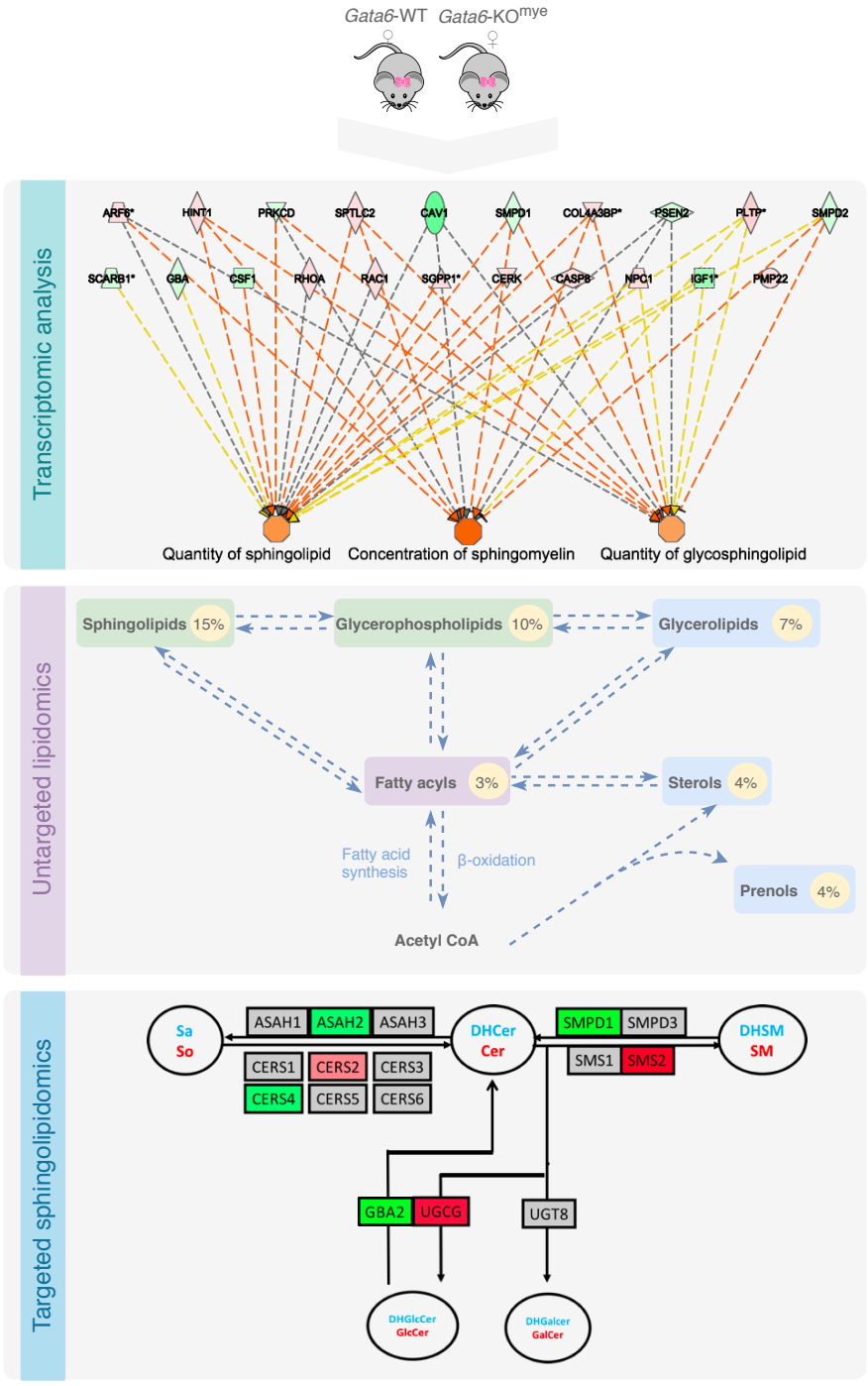


Figure 8.1 Regulation of lipid metabolism by GATA6, an integrated 'omics' approach.
The figures depicts the main findings obtained after analysing *Gata6*-KO^{mye} and WT pResMϕ from female mice, using different omics approaches. Transcriptomic analysis – GATA6 deficiency leads to activation of gene networks that suggest increased levels of SL; Untargeted lipidomics- GATA6 deficiency leads to accumulation of putatively identified SL; targeted sphingolipidomics – validation of increased SL levels and possible contribution of genes involved in their metabolism.

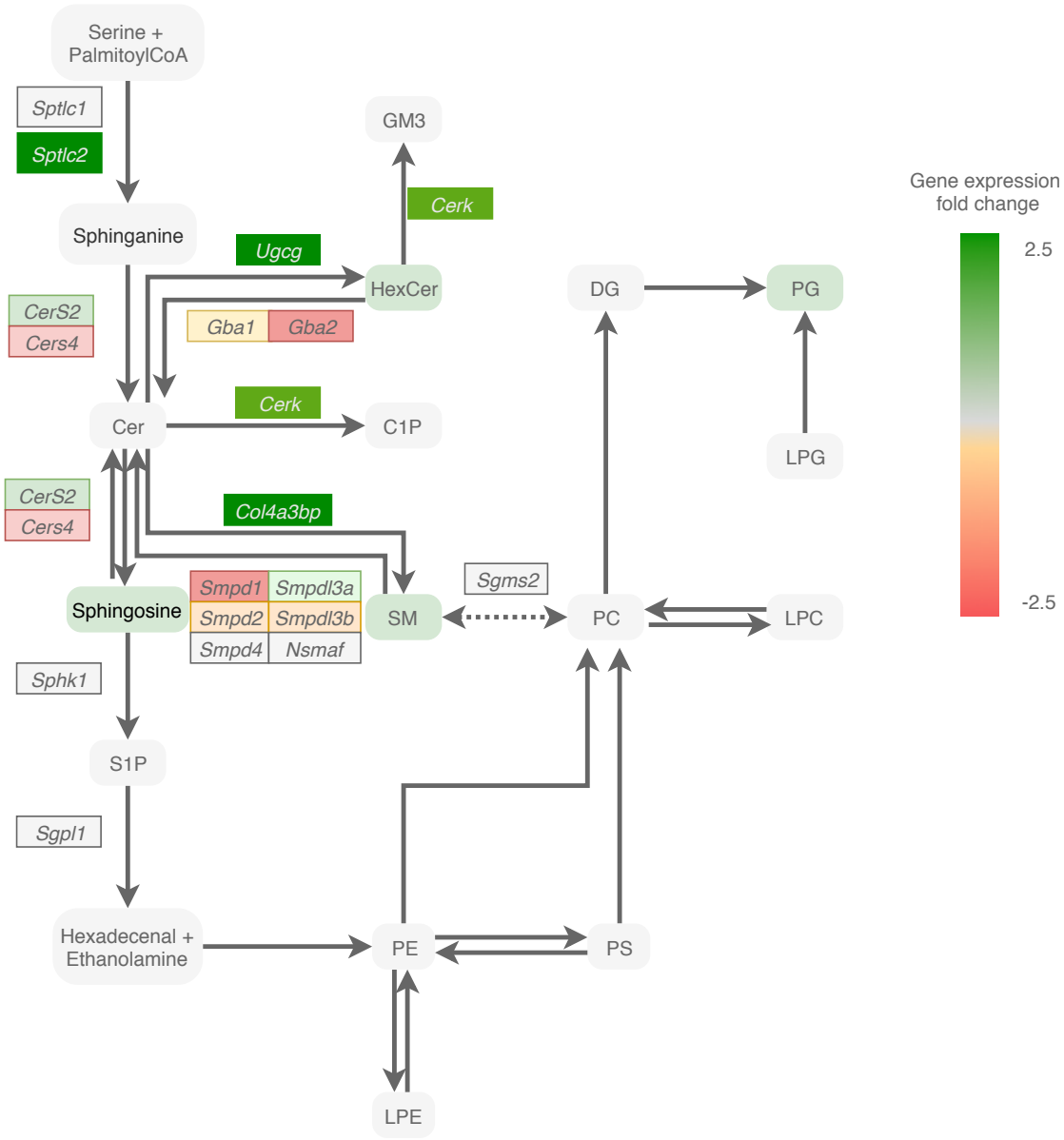


Figure 8.2 GATA6 Transcription of the SL Metabolic Network.

Selected SL and glycerophospholipid metabolic reactions (KEGG), shown together with main metabolites (rounded rectangles) and 24 selected genes (rectangles). Heatmaps show relative expression of 20 selected genes after *Gata6*-KO, measured by microarray analysis, coloured according to the scale shown on the figure. Metabolites are coloured as either green or grey depending on whether they were found to be increased, or not significantly altered, respectively.

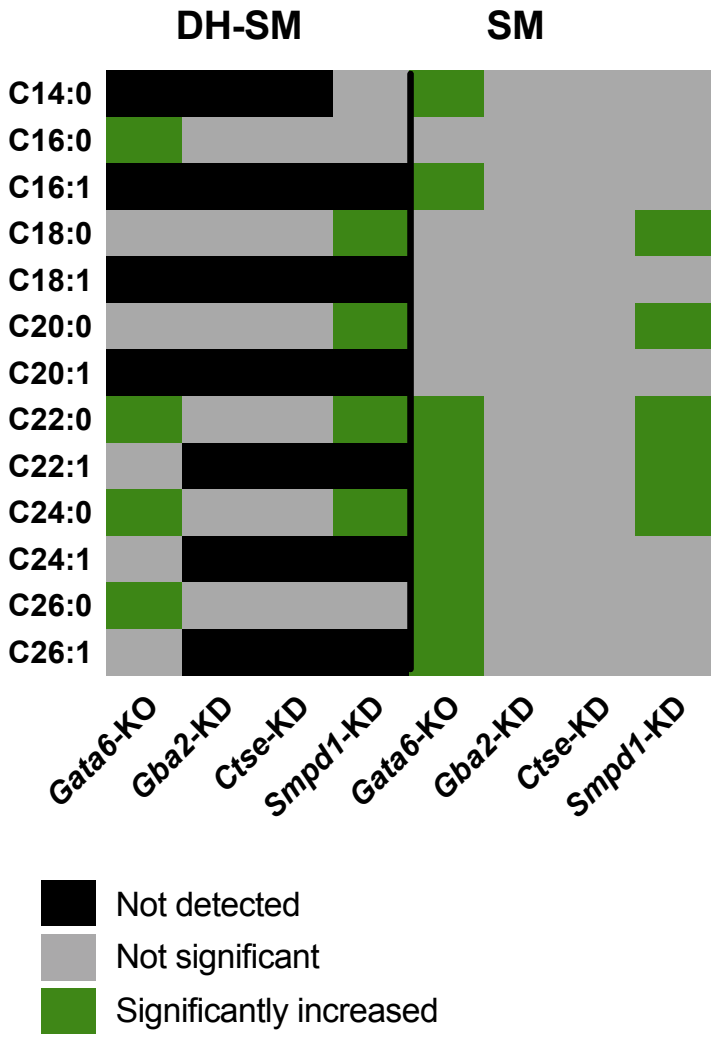


Figure 8.3 Transcriptional impact on DH-SM and SM levels in Mϕ.

Colour map describing the impact of gene expression in DH-SM and SM levels when compared to the respective controls. Each column represents a KO or KD of a specific gene, in pResMϕ (*Gata6*) and RAW 264.7 (*Gba2, Ctse, Smpd1*), whilst each row represents different chain lengths of DH-SM or SM. Black- Lipid not detected; grey- lipid detected but not significantly changed; green- lipid levels significantly increased when compare to control samples. Control for *Gata6*-KO pResMϕ were WT littermates, whilst for RAW 264.7 it was cells from the same cell line transfected with a non-silencing shRNA.

Next, I found that downregulation of both *Ctse and Gba2*, led to an increase in some of the DH-GlcCer and GlcCer species affected by GATA6 deficiency (Figure 8.4). However, the 2 knockdowns had an effect on different lipid species, with the downregulation of *Ctse* leading to an increase of most DH-GlcCer and GlcCer, whilst downregulation of GlcCer seemed to have an effect only on some chain-lengths, that matched the ones increased upon *Gata6* KO. The results obtained upon downregulation of *Ctse* were novel, as CTSE deficiency is known to lead to a lysosomal storage disorder, however, there is no information regarding how the lipidome is affected[169]. While this study does not offer a conclusive answer to the question of how the GATA6 transcription factor regulates lipid metabolism in pResMϕ, and how that impacts the cell phenotype, the findings that I have presented suggest a potential mechanism through which GATA6 controls lipid levels in pResMϕ. This is important because it contributes to our understanding of Mϕ biology and will help guide future research.

Lipids are a complex class of molecules that participate in many facets of cell structure and function. Therefore, the use of "-omic" approaches – *i.e.*, to analyse as many as possible of the molecular species and mRNAs for pathway genes – is likely to uncover interesting, and sometimes unexpected, links between their metabolism. Using the multi-faceted approach of lipidomic and genomic analyses combined techniques, my studies integrated various aspects of lipid biology in order to obtain a more clearly defined 'big picture' of the cellular effects of transcriptional regulation, on lipid metabolism and the associated functional consequences.

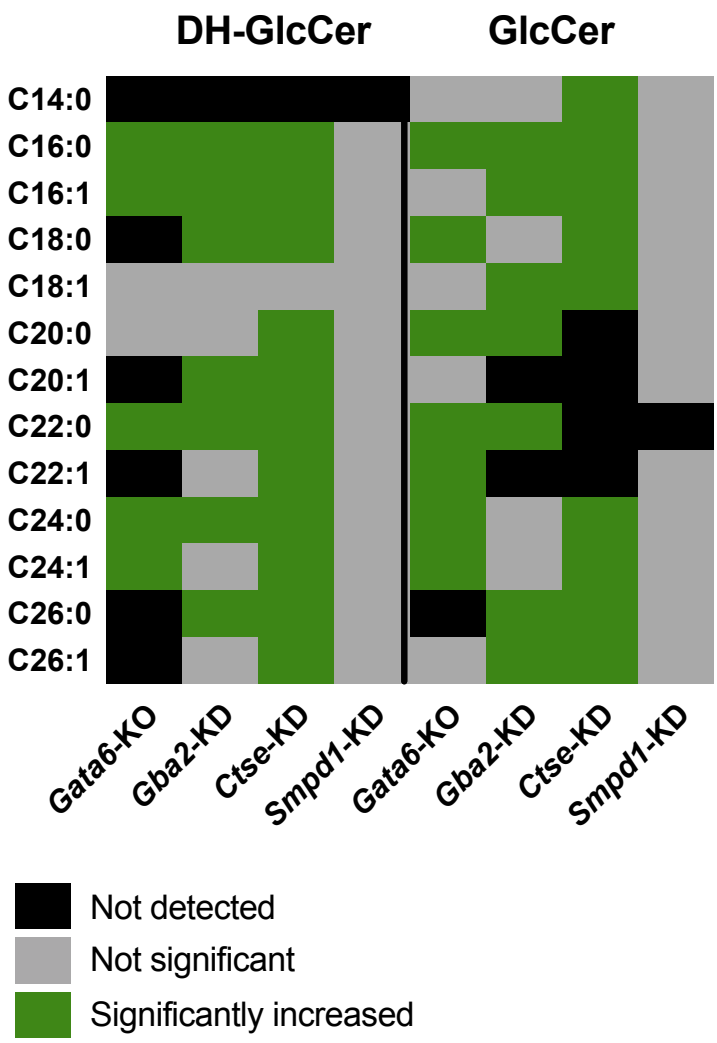


Figure 8.4 Transcriptional impact on DH-GlcCer and GlcCer levels in Mϕ.

Colour map describing the impact of gene expression in DH-GlcCer and GlcCer levels when compared to the respective controls. Each column represents a KO or KD of a specific gene, in pResMϕ (*Gata6*) and RAW 264.7 (*Gba2, Ctse, Smpd1*), whilst each row represents different chain lengths of DH-GlcCer or GlcCer.cBlack- Lipid not detected; grey- lipid detected but not significantly changed; green- lipid levels significantly increased when compare to control samples. Control for *Gata6*-KO pResMϕ were WT littermates, whilst for RAW 264.7 it was cells from the same cell line transfected with a non-silencing shRNA.

8.1 Limitations

The main limitations of this work can be categorised in into four areas. The first one lies with the untargeted lipidomics analysis performed in *Gata6*-WT and -KO^{mye} pResMϕ (chapter 5), which is an invaluable tool that allowed me to acquire an overview of the lipid changes caused by GATA6 deficiency. However, the information I can obtain from this analysis is limited by the extraction procedure, LC conditions, as well as the entries available in public lipid databases. The lipid extraction should be capable of effectively extracting lipids representative of the sample under study without bias, inducing the degradation of lipids, or introducing contamination[267]. Thus, the success in the identification and profiling of lipids is critically dependent on the efficiency of the extraction step. Herein I used an established extraction procedure known to efficiently extract different lipid categories/classes. Nevertheless, the issue still remains that some lipid species might be under-represented, which could potentially be artificially increased during the LC separation, and subsequent data analysis. Conversely, I am confident that my lipid extraction and LC-MS conditions minimised these limitations and that the conclusions drawn from the untargeted analysis are representative of the lipidome of pResMϕ.

Second, in the interest of time, and to follow a logical flow in this project, I chose to study in depth the changes cause by GATA6 deficiency in SL levels (chapter 6). In addition to SL, both FA and GPL were also affected, which can lead to significant metabolic effects in pResMϕ. Fatty acid and GPL metabolism is known to be involved in the modulation of inflammatory responses in Mϕ[268,269]. Thus, having analysed only one lipid category is a limitation, and knowing how GATA6 contributes to their regulation will help get an overview of its influence in pResMϕ phenotype.

Third, it is imperative to study the lipidomic changes in *Gata6*-WT and -KO^{mye} pResMϕ, not only in basal conditions, but also under inflammatory stimuli. This is a gap that I hope to address in the near future.

Last, the knockdown experiments in Chapter 7 were performed in RAW 264.7 cells. This was due to both time constraints and logistics, as these cells are more readily available than pResMϕ. These experiments provided valuable information, however, the fact still remains that GATA6 is not expressed in RAW 264.7 cells, and the basal levels of some lipids are significantly different between the two cell types (Chapter 4). Therefore, the mechanisms by which lipid metabolism is regulated, and its effects, might be different to pResMϕ.

In conclusion, ResMϕ represent favourable targets for modern medicine, as a wide spectrum of diseases, such as autoimmune, neurodegenerative and metabolic disorders, as well as tumour growth, have been linked to their specific dysfunctions[270]. Nevertheless, in order to be able to exploit Mϕ functionalities it is necessary to better understand the underlying mechanisms controlling their development and maintenance. I believe this thesis has contributed to this purpose. In the next section I will address the future directions, as well as how the knowledge obtained during this study can be used to modulate pResMϕ phenotype for therapeutic purposes.

8.2 Future directions

8.2.1 Inflammatory challenge

ResMϕ are tissue-specialised immune sentinels, which have key functions in homoeostasis and inflammation[2,271]. As previously mentioned, like many ResMϕ, pResMϕ are not originally derived from monocytes, but rather from embryonic progenitors seeded into tissues before birth, with the populations maintained by local proliferation[29,42,272,273]. These cells exist in complex environments and do not fit traditional polarisation categories, such as LPS and IFN-γ stimulated pro-inflammatory and IL-4 stimulated anti-inflammatory[2,274]. Furthermore, GATA6-deficient Mϕ demonstrate an altered proliferation state during peritonitis[69]. Thus, it would be relevant to investigate the changes in lipid levels under an inflammatory stimulus.

8.2.2 Gene knockdowns in pResMϕ

The majority of Mϕ research has been performed using cell lines and/or in vitro-derived BMDM. Accordingly, precise mechanisms of how cells meet metabolic demands *in vivo* remain unknown, making it critical to assess cell function in conjunction with tissue-niche metabolites[275]. Thus, in order to determine whether the modulation of *Smpd1* and *Ctse* expression by GATA6 has a significant impact in pResMϕ phenotype, the first step would be to knockdown *Smpd1* and *Ctse/Gba2* expression in WT pResMϕ.

8.2.3 Cancer research

The peritoneal cavity is a common site of tumour development and metastasis. Ovarian epithelial cancer frequently spreads to the peritoneal cavity and is treated similarly to primary peritoneal cancer[276]. There is clearly the potential for extensive cross-talk between tumour cells and pResMϕ within the peritoneal cavity. The ability for Mϕ to produce immunosuppressive cytokines, tumour-promoting growth and angiogenic factors has been well described in the literature[277]. In the peritoneum, Mϕ depletion can reduce tumour progression, metastasis[278] and ascites fluid formation[279]. Mϕ-targeted therapies are thus

promising strategies for the treatment of ovarian carcinoma and peritoneal tumours. Furthermore, GATA6 has been shown to play important roles in different types of cancer, such as, promoting cell invasion in colon cancer[280], regulating tumour dissemination in pancreatic cancer. Whilst at the same time, SL metabolism is dysregulated in cancer cells[134,281–283]. Cancer cells undergo profound changes in lipid metabolism and homeostasis, thus offering new diagnostic and therapeutic opportunities that could be unravelled by lipidomics. However, no link has been made between GATA6 and the regulation of lipid metabolism in tumour development and metastasis. Thus, I propose that future studies involving GATA6 focus on its role in cancer progression. Preliminary studies could be performed using colon cancer cell lines and characterising their phenotypes, and lipid levels upon dysregulation of targeted genes, such as the transcription factor GATA6, as well as *Smpd1*, *Gba1/Gba2* and, *Ctse* , which I have found to be key players in the regulation of lipid metabolism in pResMϕ.

8.2.4 Lipidomics analysis

Lastly, the lipidomics studies performed during the course of this project were based on an untargeted lipidomics approach, used to obtain an overview of the impact of GATA6-deficiency on lipid metabolism, followed by a targeted approach applied to the SL category. However, an in-depth analysis of the lipidomic changes caused by GATA6-defficiency on other lipid categories, such as FA and GPL, would contribute to the knowledge of how GATA6 regulates lipid metabolism in pResMϕ. Additionally, detailed in-sight into lipid function requires knowledge of the subcellular localisation of individual lipids. Hence, it would be of value to perform fractional organelle lipidomics on *Gata6*-WT and KO^{mye} pResMϕ.

References

1. Gordon, S. Elie Metchnikoff: Father of natural immunity. *Eur. J. Immunol.* **38,** 3257–3264 (2008).

2. Davies, L. C. & Taylor, P. R. Tissue-resident macrophages: then and now. *Immunology* **144,** 541–548 (2015).

3. Hoeffel, G. & Ginhoux, F. Fetal monocytes and the origins of tissue-resident macrophages. *Cell. Immunol.* (2018).

4. Stefater, J. A., Ren, S., Lang, R. A. & Duffield, J. S. Metchnikoff's policemen: macrophages in development, homeostasis and regeneration. *Trends Mol. Med.* **17,** 743–752 (2011).

5. van Furth, R. *et al.* The mononuclear phagocyte system: a new classification of macrophages, monocytes, and their precursor cells. *Bull. World Health Organ.* **46,** 845–852 (1972).

6. Gordon, S. The macrophage: Past, present and future. *Eur. J. Immunol.* **37,** 9–17 (2007).

7. Hume, D. A. The mononuclear phagocyte system. *Curr. Opin. Immunol.* **18,** 49–53 (2006).

8. Gordon, S. & Plüddemann, A. Tissue macrophages: Heterogeneity and functions. *BMC Biol.* **15,** 1–18 (2017).

9. Jones, C. V & Ricardo, S. D. Macrophages and CSF-1: implications for development and beyond. *Organogenesis* **9,** 249–60 (2013).

10. Mass, E. *et al.* Specification of tissue-resident macrophages during organogenesis. *Science (80-.).* **353,** aaf4238-aaf4238 (2016).

11. Eom, D. S. & Parichy, D. M. A macrophage relay for long-distance signaling during postembryonic tissue remodeling. *Science (80-.).* **355,** 1317–1320 (2017).

12. DeFalco, T., Bhattacharya, I., Williams, A. V., Sams, D. M. & Capel, B. Yolk-sac-derived macrophages regulate fetal testis vascularization and morphogenesis. *Proc. Natl. Acad. Sci.* **111,** E2384–E2393 (2014).

13. Keller, N. M. *et al.* Developmental origin and maintenance of distinct testicular macrophage populations. *Jem* **214,** 1–14 (2017).

14. Rae, F. *et al.* Characterisation and trophic functions of murine embryonic macrophages based upon the use of a Csf1r-EGFP transgene reporter. *Dev. Biol.* **308,** 232–246 (2007).

15. Liu, J. *et al.* CCR2-and CCR2+corneal macrophages exhibit distinct characteristics and balance inflammatory responses after epithelial abrasion. *Mucosal Immunol.* **10,** 1145–1159 (2017).

16. Reyes, N. J., O'Koren, E. G. & Saban, D. R. New insights into mononuclear phagocyte biology from the visual system. *Nat. Rev. Immunol.* **17,** 322–332 (2017).

17. McGrath, K. E. *et al.* Enucleation of primitive erythroid cells generates a transient population of 'pyrenocytes' in the mammalian fetus. *Blood* **111,** 2409–2417 (2008).

18. Lavine, K. J. *et al.* Distinct macrophage lineages contribute to disparate patterns of cardiac recovery and remodeling in the neonatal and adult heart. *Proc. Natl. Acad. Sci. U. S. A.* **111,** (2014).

19. Wynn, T. A. & Vannella, K. M. Macrophages in Tissue Repair, Regeneration, and Fibrosis. *Immunity* **44,** 450–462 (2016).

20. Epelman, S., Lavine, K. J. & Randolph, G. J. Origin and Functions of Tissue Macrophages. *Immunity* **41,** 21–35 (2014).

21. Geissmann, F. *et al.* Development of monocytes, macrophages, and dendritic cells. *Science* **327,** 656–661 (2010).

22. Pittet, M. J., Nahrendorf, M. & Swirski, F. K. The journey from stem cell to macrophage. *Ann. N. Y. Acad. Sci.* **1319,** 1–18 (2014).

23. Zhao, Y., Zou, W., Du, J. & Zhao, Y. The origins and homeostasis of monocytes and tissue-resident macrophages in physiological situation. *J. Cell. Physiol.* (2018).

24. Davies, L. C. *et al.* Distinct bone marrow-derived and tissue-resident macrophage lineages proliferate at key stages during inflammation. *Nat. Commun.* **4,** 1886 (2013).

25. Sieweke, M. H. & Allen, J. E. Beyond stem cells: self-renewal of differentiated macrophages. *Science* **342,** 1242974 (2013).

26. Gundra, U. M. *et al.* Alternatively activated macrophages derived from monocytes and tissue macrophages are phenotypically and functionally distinct. *Blood* **123,** e110-22 (2014).

27. Davies, L. C., Jenkins, S. J., Allen, J. E. & Taylor, P. R. Tissue-resident macrophages. *Nat. Immunol.* **14,** 986–95 (2013).

28. Perdiguero, E. G. *et al.* Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature* (2014).

29. Yona, S. *et al.* Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity* **38,** 79–91 (2013).

30. Shi, C. & Pamer, E. G. Monocyte recruitment during infection and inflammation. *Nat. Rev. Immunol.* **11,** 762–74 (2011).

31. Ingersoll, M. a, Platt, A. M., Potteaux, S. & Randolph, G. J. Monocyte trafficking in acute and chronic inflammation. *Trends Immunol.* **32,** 470–7 (2011).

32. Mogensen, T. H. Pathogen recognition and inflammatory signaling in innate immune defenses. *Clin. Microbiol. Rev.* **22,** 240–73, Table of Contents (2009).

33. Newton, K. & Dixit, V. M. Signaling in Innate Immunity and Inflammation. *Cold Spring Harb. Perspect. Biol.* **4,** a006049–a006049 (2012).

34. Crocker, P. R. & Gordon, S. Isolation and characterization of resident stromal macrophages and hematopoietic cell clusters from mouse bone marrow. *J. Exp. Med.* **162,** 993–1014 (1985).

35. Dantzer, R., O'Connor, J. C., Freund, G. G., Johnson, R. W. & Kelley, K. W. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat. Rev. Neurosci.* **9,** 46–56 (2008).

36. Wynn, T. A., Chawla, A. & Pollard, J. W. Macrophage biology in development, homeostasis and disease. *Nature* **496,** 445–455 (2013).

37. Gustafson, B., Hedjazifar, S., Gogg, S., Hammarstedt, A. & Smith, U. Insulin resistance and impaired adipogenesis. *Trends Endocrinol. Metab.* **26,** 193–200 (2015).

38. Boutens, L. & Stienstra, R. Adipose tissue macrophages: going off track during obesity. *Diabetologia* **59,** 879–94 (2016).

39. Liddelow, S. A. *et al.* Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* **541,** 481–487 (2017).

40. Mass, E. Delineating the origins, developmental programs and homeostatic functions of tissue-resident macrophages. *Int. Immunol.* (2018).

41. Haldar, M. & Murphy, K. M. Origin, development, and homeostasis of tissue-resident macrophages. *Immunol. Rev.* **262,** 25–35 (2014).

42. Bain, C. C. *et al.* Long-lived self-renewing bone marrow-derived macrophages displace embryo-derived cells to inhabit adult serous cavities. *Nat. Commun.* **7,** 1–14 (2016).

43. Hashimoto, D. *et al.* Adult Life With Minimal Contribution From Circulating Monocytes. *J.Immuni* **38,** 1–25 (2014).

44. Tamoutounour, S. *et al.* Origins and functional specialization of macrophages and of conventional and monocyte-derived dendritic cells in mouse skin. *Immunity* **39,** 925–938 (2013).

45. Bain, C. C. *et al.* Constant replenishment from circulating monocytes maintains the macrophage pool in the intestine of adult mice. *Nat. Immunol.* **15,** 929–937 (2014).

46. Epelman, S. *et al.* Embryonic and Adult-Derived Resident Cardiac Macrophages Are Maintained through Distinct Mechanisms at Steady State and during Inflammation. *Immunity* **40,** 91–104 (2014).

47. Molawi, K. *et al.* Progressive replacement of embryo-derived cardiac macrophages with age. *J. Exp. Med.* **211,** 2151–2158 (2014).

48. Ghigo, C. *et al.* Multicolor fate mapping of Langerhans cell homeostasis. *J. Exp. Med.* **210,** 1657–1664 (2013).

49. Gomez Perdiguero, E. *et al.* Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature* **518,** 547–551 (2015).

50. Hoeffel, G. *et al.* Adult Langerhans cells derive predominantly from embryonic fetal liver monocytes with a minor contribution of yolk sac-derived macrophages. *J. Exp. Med.* **209,** 1167–81 (2012).

51. Okabe, Y. & Medzhitov, R. Tissue-specific signals control reversible program of localization and functional polarization of macrophages. *Cell* **157,** 832–44 (2014).

52. Lavin, Y. *et al.* Tissue-Resident Macrophage Enhancer Landscapes Are Shaped by the Local Microenvironment. *Cell* **159,** 1312–1326 (2014).

53. Gosselin, D. *et al.* Environment drives selection and function of enhancers controlling tissue-specific macrophage identities. *Cell* **159,** 1327–1340 (2014).

54. Gordon, S. & Taylor, P. R. Monocyte and macrophage heterogeneity. *Nat. Rev. Immunol.* **5,** 953–64 (2005).

55. Murray, P. J. & Wynn, T. A. Protective and pathogenic functions of macrophage subsets. *Nat. Rev. Immunol.* **11,** 723–737 (2011).

56. William J. Boyle, W. Scott Simonet & David L. Lacey. Osteoclast differentiation and activation. *Nature* **423,** 337–342 (2003).

57. Gautier, E. L. *et al.* Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nat. Immunol.* **13,** 1118–28 (2012).

58. Kohyama, M. *et al.* Role for Spi-C in the development of red pulp macrophages and splenic iron homeostasis. *Nature* **457,** 318–321 (2009).

59. A-Gonzalez, N. *et al.* The nuclear receptor LXRα controls the functional specialization of splenic macrophages. *Nat. Immunol.* **14,** 831–839 (2013).

60. Takayanagi, H. *et al.* Induction and activation of the transcription factor NFATc1 (NFAT2) integrate RANKL signaling in terminal differentiation of osteoclasts. *Dev. Cell* **3,** 889–901 (2002).

61. Schulz, C. *et al.* A Lineage of Myeloid Cells Independent of Myb and Hematopoietic Stem Cells. *Science (80-.).* **336,** 86–90 (2012).

62. Gautier, E. L. *et al.* Gene expression profiles and transcriptional regulatory pathways underlying mouse tissue macrophage identity and diversity. *Nat Immunol* **13,** 1118–1128 (2012).

63. Rosas, M., Thomas, B., Stacey, M., Gordon, S. & Taylor, P. R. The myeloid 7/4-antigen defines recently generated inflammatory macrophages and is synonymous with Ly-6B. *J Leukoc Biol* **88,** 169–180 (2010).

64. Taylor, P. R., Brown, G. D., Geldhof, A. B., Martinez-Pomares, L. & Gordon, S. Pattern recognition receptors and differentiation antigens define murine myeloid cell heterogeneity ex vivo. *Eur. J. Immunol.* **33,** 2090–2097 (2003).

65. Cain, D. W. *et al.* Identification of a Tissue-Specific, C/EBP -Dependent Pathway of Differentiation for Murine Peritoneal Macrophages. *J. Immunol.* **191,** 4665–4675 (2013).

66. Cassado, Alexandra dos Anjos Lima, M. R. & Bortoluci, K. R. Revisiting Mouse Peritoneal Macrophages: Heterogeneity, Development, and Function. *Front. Immunol.* **6,** 225 (2015).

67. Cesar Machado, M. C. & Mendonca Coelho, A. M. in *Acute Pancreatitis* (InTech, 2012).

68. Dahdah, A. *et al.* Mast cells aggravate sepsis by inhibiting peritoneal macrophage phagocytosis. *J. Clin. Invest.* **124,** 4577–4589 (2014).

69. Rosas, M. *et al.* The transcription factor Gata6 links tissue macrophage phenotype and proliferative renewal. *Science* **344,** 645–648 (2014).

70. Wang, C. *et al.* Characterization of murine macrophages from bone marrow, spleen and peritoneum. *BMC Immunol.* **14,** 1–10 (2013).

71. Gautier, E. L. *et al.* Gata6 regulates aspartoacylase expression in resident peritoneal macrophages and controls their survival. *J. Exp. Med.* **211,** 1525–31 (2014).

72. Viger, R. S., Guittot, S. M., Anttonen, M., Wilson, D. B. & Heikinheimo, M. Role of the GATA Family of Transcription Factors in Endocrine Development, Function, and Disease. *Mol. Endocrinol.* **22,** 781–798 (2008).

73. Ko, L. J. & Engel, J. D. DNA-binding specificities of the GATA transcription factor family. *Mol. Cell. Biol.* **13,** 4011–22 (1993).

74. Pardo, M. *et al.* An Expanded Oct4 Interaction Network: Implications for

Stem Cell Biology, Development, and Disease. *Cell Stem Cell* **6,** 382–395 (2010).

75. van den Berg, D. L. C. *et al.* An Oct4-Centered Protein Interaction Network in Embryonic Stem Cells. *Cell Stem Cell* **6,** 369–381 (2010).

76. Shu, J. *et al.* GATA family members as inducers for cellular reprogramming to pluripotency. *Cell Res.* **25,** 169–180 (2015).

77. Molkentin, J. D. The zinc finger-containing transcription factors GATA-4, -5, and -6: Ubiquitously expressed regulators of tissue-specific gene expression. *J. Biol. Chem.* **275,** 38949–38952 (2000).

78. Morrisey, E. E., Ip, H. S., Lu, M. M. & Parmacek, M. S. GATA-6: a zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm. *Dev. Biol.* **177,** 309–22 (1996).

79. Maeda, M., Kubo, K., Nishi, T. & Futai, M. Roles of gastric GATA DNA-binding proteins. *J Exp Biol* **199,** 513–520 (1996).

80. Huggon, I. C. *et al.* Molecular cloning of human GATA-6 DNA binding protein: High levels of expression in heart and gut. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1353,** 98–102 (1997).

81. Takeda, M., Obayashi, K., Kobayashi, A. & Maeda, M. A unique role of an amino terminal 16-residue region of long-type GATA-6. *J. Biochem.* **135,** 639–650 (2004).

82. Brewer, A. *et al.* The Human and Mouse GATA-6 Genes Utilize Two Promoters and Two Initiation Codons. *J. Biol. Chem.* **274,** 38004–38016 (1999).

83. Rogers, S., Wells, R. & Rechsteiner, M. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science (80-.).* **234,** 364–368 (1986).

84. Suzuki, E. *et al.* The HumanGATA-6Gene: Structure, Chromosomal Location, and Regulation of Expression by Tissue-Specific and Mitogen-Responsive Signals. *Genomics* **38,** 283–290 (1996).

85. Lavin, Y., Mortha, A., Rahman, A. & Merad, M. Regulation of macrophage development and function in peripheral tissues. *Nat. Rev. Immunol.* **15,** 731–44 (2015).

86. Karasinska, J. M. *et al.* ABCA1 influences neuroinflammation and neuronal death. *Neurobiol. Dis.* **54,** 445–455 (2013).

87. Lin, Y. C., Lin, C. H., Kuo, C. Y. & Yang, V. C. ABCA1 modulates the oligomerization and Golgi exit of caveolin-1 during HDL-mediated cholesterol efflux in aortic endothelial cells. *Biochem. Biophys. Res. Commun.* **382,** 189–195 (2009).

88. Kuhn, H., Banthiya, S., Leyen, K. Van & General, M. HHS Public Access. **1851,** 308–330 (2015).

89. Dennis, E. A. Lipidomics joins the omics evolution. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 2089–90 (2009).

90. Papin, J. A., Hunter, T., Palsson, B. O., Subramaniam, S. & Pines, N. T. Reconstruction of cellular signalling networks and analysis of their properties. **6,** (2005).

91. Diseases, I. Overview of the Alliance for Cellular Signaling. 703–706 (2002).

92. Li Feng, G. D. P. *Functional lipidomics*. (Taylor & Francis Group, 2005).

93. Slatter, D. A. *et al.* Mapping the Human Platelet Lipidome Reveals Cytosolic Phospholipase A 2 as a Regulator of Mitochondrial Bioenergetics during Activation Resource Mapping the Human Platelet Lipidome Reveals Cytosolic Phospholipase A 2 as a Regulator of Mitochondrial Bioene. *Cell Metab.* **23,** 930–944 (2016).

94. Koelmel, J. P. *et al.* Expanding Lipidome Coverage Using LC-MS/MS Data-Dependent Acquisition with Automated Exclusion List Generation. *J. Am. Soc. Mass Spectrom.* **28,** 908–917 (2017).

95. Dennis, E. A. *et al.* A mouse macrophage lipidome. *J. Biol. Chem.* **285,** 39976–85 (2010).

96. Quehenberger, O. & Dennis, E. A. The human plasma lipidome. *N. Engl. J. Med.* **365,** 1812–23 (2011).

97. Han, X. *Lipidomics - Comprehensive mass spectrometry of lipids*. (John Wiley & Sons, Inc, 2016).

98. Köberlin, M. S. S. *et al.* A Conserved Circular Network of Coregulated Lipids Modulates Innate Immune Responses. *Cell* **162,** 1–14 (2015).

99. Fahy, E. *et al.* A comprehensive classification system for lipids 1. **46,** (2005).

100. Hadadi, N. *et al.* A computational framework for integration of lipidomics data into metabolic pathways. *Metab. Eng.* **23,** 1–8 (2014).

101. Gross, R. W. & Han, X. Lipidomics at the interface of structure and function in systems biology. *Chem. Biol.* **18,** 284–291 (2011).

102. Murphy, R. C., Fiedler, J. & Hevko, J. Analysis of nonvolatile lipids by mass spectrometry. *Chem. Rev.* **101,** 479–526 (2001).

103. Ivanova, P. T. *et al.* Electrospray ionization mass spectrometry analysis of changes in phospholipids in RBL-2H3 mastocytoma cells during degranulation. (2001).

104. Wenk, M. R. The emerging field of lipidomics. *Nat. Rev. Drug Discov.* **4,** 594–610 (2005).

105. Meer, G. Van. New EMBO Member ' s Review Cellular lipidomics. **24,** 3159–3165 (2005).

106. Sud, M. *et al.* LMSD: LIPID MAPS structure database. *Nucleic Acids Res.* **35,** 527–532 (2007).

107. Alhouayek, M. & Muccioli, G. G. in *Handbook of Cannabis and Related Pathologies* 616–629 (Elsevier, 2017). doi:10.1016/B978-0-12-800756-3.00074-0

108. Howard, J. G., Christie, G. H., Boakt, J. L. & Kinskyt, R. G. Peritoneal and alveolar macrophages derived from lymphocyte populations during graft-versus-host reaction. *Br. J. Exp. Pathol.* (1969).

109. Fernandis, A. Z. & Wenk, M. R. Membrane lipids as signaling molecules. *Curr. Opin. Lipidol.* **18,** 121–8 (2007).

110. Huber, L. C. *et al.* The role of membrane lipids in the induction of macrophage apoptosis by microparticles. *Apoptosis* **12,** 363–374 (2007).

111. Norris, P. C. & Dennis, E. a. A lipidomic perspective on inflammatory macrophage eicosanoid signaling. *Adv. Biol. Regul.* **54,** 99–110 (2014).

112. Shimizu, T. Lipid Mediators in Health and Disease: Enzymes and Receptors as Therapeutic Targets for the Regulation of Immunity and Inflammation. *Annu. Rev. Pharmacol. Toxicol.* **49,** 123–150 (2009).

113. Serhan, C. N. & Savill, J. Resolution of inflammation: the beginning programs the end. *Nat. Immunol.* **6,** 1191–1197 (2005).

114. El Alwani, M., Wu, B. X., Obeid, L. M. & Hannun, Y. A. Bioactive sphingolipids in the modulation of the inflammatory response. *Pharmacol. Ther.* **112,** 171–183 (2006).

115. Smith, E. R., Merrill, A. H., Obeid, L. M. & Hannun, Y. A. Effects of sphingosine and other sphingolipids on protein kinase C. *Methods Enzymol.* **312,** 361–73 (2000).

116. Venable, M. E., Lee, J. Y., Smyth, M. J., Bielawska, A. & Obeid, L. M. Role of Ceramide in Cellular Senescence. *J. Biol. Chem.* **270,** 30701–30708 (1995).

117. Obeid, L. M., Linardic, C. M., Karolak, L. a & Hannun, Y. a. Programmed cell death induced by ceramide. *Science* **259,** 1769–1771 (1993).

118. Hannun, Y. A. & Obeid, L. M. Principles of bioactive lipid signalling: lessons from sphingolipids. *Nat. Rev. Mol. Cell Biol.* **9,** 139–150 (2008).

119. Chalfant, C. E. Sphingosine 1-phosphate and ceramide 1-phosphate: expanding roles in cell signaling. *J. Cell Sci.* **118,** 4605–4612 (2005).

120. Mitsutake, S. *et al.* Ceramide Kinase Is a Mediator of Calcium-dependent Degranulation in Mast Cells. *J. Biol. Chem.* **279,** 17570–17577 (2004).

121. Thudichum, J. L. W. (John L. W. *A treatise on the chemical constitution of the brain : based throughout upon original researches*. (London : Bailliere, Tindall, and Cox, 1884).

122. Testi, R. Sphingomyelin breakdown and cell fate. *Trends Biochem. Sci.* **21,** 468–471 (1996).

123. Chalfant, C. & Poeta, M. Del. *Sphingolipids as signaling and regulatory molecules*.

124. Futerman, A. H. & Hannun, Y. A. The complex life of simple sphingolipids. *EMBO Rep.* **5,** 777–782 (2004).

125. Hla, T. & Dannenberg, A. J. Sphingolipid signaling in metabolic disorders. *Cell Metab.* **16,** 420–434 (2012).

126. Merrill, A. H. & Carman, G. M. Introduction to thematic minireview series: Novel bioactive sphingolipids. *J. Biol. Chem.* **290,** 15362–15364 (2015).

127. Merrill, A. H., Wang, M. D., Park, M. & Sullards, M. C. (Glyco)sphingolipidology: an amazing challenge and opportunity for systems biology. *Trends Biochem. Sci.* **32,** 457–468 (2007).

128. Merrill, A. H. Sphingolipid and glycosphingolipid metabolic pathways in the era of sphingolipidomics. *Chem. Rev.* **111,** 6387–6422 (2011).

129. Don, A. S., Lim, X. Y. & Couttas, T. A. Re-configuration of sphingolipid metabolism by oncogenic transformation. *Biomolecules* **4,** 315–353 (2014).

130. Gault, C. R., Obeid, L. M. & Hannun, Y. A. An overview of sphingolipid metabolism: From synthesis to breakdown. *Adv. Exp. Med. Biol.* **688,** 1–23 (2010).

131. Sims, K. H. Participation of de de novo sphingolipid biosynthesis in the regulation of autophagy in response to diverse agents. (2011).

132. Here, A. & Homology, T. Molecular machinery for non-vesicular traf cking of ceramide. *Nature* **426,** 1–7 (2003).

133. Mandon, E. C., Ehses, I., Rother, J., Van Echten, G. & Sandhoff, K. Subcellular localization and membrane topology of serine palmitoyltransferase, 3-dehydrosphinganine reductase, and sphinganine N- acyltransferase in mouse liver. *J. Biol. Chem.* **267,** 11144–11148 (1992).

134. Zheng, W. *et al.* Ceramides and other bioactive sphingolipid backbones in health and disease: Lipidomic analysis, metabolism and roles in membrane structure, dynamics, signaling and autophagy. *Biochim. Biophys. Acta - Biomembr.* **1758,** 1864–1884 (2006).

135. Pewzner-Jung, Y., Ben-Dor, S. & Futerman, A. H. When do Lasses (longevity assurance genes) become CerS (ceramide synthases)? Insights into the regulation of ceramide synthesis. *J. Biol. Chem.* **281,** 25001–25005 (2006).

136. Michel, C. *et al.* Characterization of Ceramide Synthesis. *J. Biol. Chem.* **272,** 22432–22437 (1997).

137. Merrill, A. H. & Wang, E. Biosynthesis of long-chain (sphingoid) bases from serine by LM cells. Evidence for introduction of the 4-trans-double bond after de novo biosynthesis of N-acylsphinganine(s). *J. Biol. Chem.* **261,** 3764–3769 (1986).

138. Venkataraman, K. & Futerman, A. H. Ceramide as a second messenger: Sticky solutions to sticky problems. *Trends Cell Biol.* **10,** 408–412 (2000).

139. Koch, J. *et al.* Molecular Cloning and Characterization of a Full-length Complementary DNA Encoding Human Acid Ceramidase. *J. Biol. Chem.* **271,** 33110–33115 (1996).

140. Mao, C. *et al.* Cloning and characterization of a novel human alkaline ceramidase: A mammalian enzyme that hydrolyzes phytoceramide. *J. Biol. Chem.* **276,** 26577–26588 (2001).

141. Hait, N. C., Oskeritzian, C. A., Paugh, S. W., Milstien, S. & Spiegel, S. Sphingosine kinases, sphingosine 1-phosphate, apoptosis and diseases. *Biochim. Biophys. Acta - Biomembr.* **1758,** 2016–2026 (2006).

142. Reiss, U. *et al.* Sphingosine-phosphate Lyase Enhances Stress-induced Ceramide Generation and Apoptosis. *J. Biol. Chem.* **279,** 1281–1290 (2004).

143. Loizides-Mangold, U. On the future of mass-spectrometry-based lipidomics. *FEBS J.* **280,** 2817–2829 (2013).

144. Zubarev, R. a. & Makarov, A. Orbitrap mass spectrometry. *Anal. Chem.* **85,** 5288–5296 (2013).

145. Banerjee, S. & Mazumdar, S. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *Int. J. Anal. Chem.* **2012,** 282574 (2012).

146. Nair, H. & Clarke, W. *Mass spectrometry for the clinical laboratory.* (2017). doi:10.1016/B978-0-12-800871-3/00014-6

147. Murphy, R. C. *Tandem mass spectrometry of lipids*. **179,** (The Royal Society of Chemistry, 2015).

148. Vaz, F. M., Pras-Raves, M., Bootsma, A. H. & van Kampen, A. H. C. Principles and practice of lipidomics. *J. Inherit. Metab. Dis.* (2014).

149. Erve, J. C. L., Gu, M., Wang, Y., DeMaio, W. & Talaat, R. E. Spectral Accuracy of Molecular Ions in an LTQ/Orbitrap Mass Spectrometer and Implications for Elemental Composition Determination. *J. Am. Soc. Mass Spectrom.* **20,** 2058–2069 (2009).

150. Makarov, A. *et al.* Performance Evaluation of a Hybrid Linear Ion Trap / Orbitrap Mass Spectrometer. **78,** 2113–2120 (2006).

151. Eliuk, S. & Makarov, A. Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annu. Rev. Anal. Chem.* **8,** 61–80 (2015).

152. Ståhlman, M. *et al.* High-throughput shotgun lipidomics by quadrupole time-of-flight mass spectrometry. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **877,** 2664–2672 (2009).

153. Harkewicz, R. & Dennis, E. A. Applications of Mass Spectrometry to Lipids and Membranes. *Annu. Rev. Biochem.* **80,** 301–325 (2011).

154. Ivanova, P. T., Milne, S. B., Myers, D. S. & Brown, H. A. Lipidomics : a mass spectrometry based systems level analysis of cellular lipids. (2009).

155. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods* **25,** 402–408 (2001).

156. Shaner, R. L. *et al.* Quantitative analysis of sphingolipids for lipidomics using triple quadrupole and quadrupole linear ion trap mass spectrometers. *J. Lipid Res.* **50,** 1692–1707 (2009).

157. Kessner, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. ProteoWizard: Open source software for rapid proteomics tools development. *Bioinformatics* **24,** 2534–2536 (2008).

158. Dunn, W. B. *et al.* Metabolic profiling of serum using Ultra Performance Liquid Chromatography and the LTQ-Orbitrap mass spectrometry system. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **871,** 288–298 (2008).

159. Leavy, O. Macrophages: Peritoneal population depends on GATA6. *Nat. Rev. Immunol.* **14,** 360 (2014).

160. Dempster, A. P. An overview of multivariate data analysis. *J. Multivar. Anal.* **1,** 316–346 (1971).

161. Berridge, M. J. & Irvine, R. F. Inositol phosphates and cell signalling. *Nature* **341,** 197–205 (1989).

162. Smith, W. L., Urade, Y. & Jakobsson, P.-J. Enzymes of the Cyclooxygenase Pathways of Prostanoid Biosynthesis. *Chem. Rev.* **111,** 5821–5865 (2011).

163. Doherty, G. J. & McMahon, H. T. Mechanisms of Endocytosis. *Annu. Rev. Biochem.* **78,** 857–902 (2009).

164. Le Roy, C. & Wrana, J. L. Clathrin- and non-clathrin-mediated endocytic regulation of cell signalling. *Nat. Rev. Mol. Cell Biol.* **6,** 112–26 (2005).

165. Tsukubas, T. Isolation and Characterization of Recombinant Human Cathepsin E Expressed in Chinese Hamster Ovary Cells *. **1151,** 7276–7282 (1993).

166. Gotthardt, D. *et al.* High-resolution dissection of phagosome maturation reveals distinct membrane trafficking phases. *Mol. Biol. Cell* **13,** 3508–20 (2002).

167. Krämer, A., Green, J., Pollard, J. & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30,** 523–530 (2014).

168. Sastradipura, D. F. *et al.* Identification of Cellular Compartments Involved in Processing of Cathepsin E in Primary Cultures of Rat Microglia. *J. Neurochem.* **70,** 2045–2056 (2002).

169. Yanagawa, M. *et al.* Cathepsin E deficiency induces a novel form of lysosomal storage disorder showing the accumulation of lysosomal membrane sialoglycoproteins and the elevation of lysosomal pH in macrophages. *J. Biol. Chem.* **282,** 1851–1862 (2007).

170. Kadowaki, T. *et al.* Defective adipose tissue development associated with hepatomegaly in cathepsin E-deficient mice fed a high-fat diet. *Biochem. Biophys. Res. Commun.* **446,** (2014).

171. Sonnino, S. & Prinetti, A. Sphingolipids and membrane environments for caveolin. *FEBS Lett.* **583,** 597–606 (2009).

172. Kiss, A. L. & Botos, E. Endocytosis via caveolae: Alternative pathway with distinct cellular compartments to avoid lysosomal degradation? *J. Cell. Mol. Med.* **13,** 1228–1237 (2009).

173. Fessler, M. B. & Parks, J. S. Intracellular Lipid Flux and Membrane Microdomains as Organizing Principles in Inflammatory Cell Signaling. *J. Immunol.* **187,** 1529–1535 (2011).

174. Gargalovic, P. & Dory, L. Caveolins and macrophage lipid metabolism. *J. Lipid Res.* **44,** 11–21 (2003).

175. Maceyka, M. & Spiegel, S. Sphingolipid metabolites in inflammatory disease. *Nature* **510,** 58–67 (2014).

176. Han, X. *et al.* Metabolomics in early Alzheimer's disease: Identification of altered plasma sphingolipidome using shotgun lipidomics. *PLoS One* **6,** (2011).

177. Fhaner, C. J., Liu, S., Ji, H., Simpson, R. J. & Reid, G. E. Comprehensive Lipidome Profiling of Isogenic Primary and Metastatic Colon Adrenocarcinoma Cell Lines. *Anal. Chem.* **84,** 8917–8926 (2012).

178. Cheng, H., Wang, M., Li, J. L., Cairns, N. J. & Han, X. Specific changes of sulfatide levels in individuals with pre-clinical Alzheimer's disease: An early event in disease pathogenesis. *J. Neurochem.* **127,** 733–738 (2013).

179. Qiu, B. *et al.* HIF2α-dependent lipid storage promotes endoplasmic reticulum homeostasis in clear-cell renal cell carcinoma. *Cancer Discov.* **5,** 653–667 (2016).

180. Hopperton, K. E., Duncan, R. E., Bazinet, R. P. & Archer, M. C. Fatty acid synthase plays a role in cancer metabolism beyond providing fatty acids for phospholipid synthesis or sustaining elevations in glycolytic activity. *Exp. Cell Res.* **320,** 302–310 (2014).

181. Shui, G. *et al.* Comparative plasma lipidome between human and cynomolgus monkey: Are plasma polar lipids good biomarkers for diabetic monkeys? *PLoS One* **6,** e19731 (2011).

182. Merrill, A. H., Dennis, E. A., McDonald, J. G. & Fahy, E. Lipidomics Technologies at the End of the First Decade and the Beginning of the Next. *Adv. Nutr. An Int. Rev. J.* **4,** 565–567 (2013).

183. Tautenhahn, R., Patti, G. J., Rinehart, D. & Siuzdak, G. XCMS online: A web-based platform to process untargeted metabolomic data. *Anal. Chem.* **84,** 5035–5039 (2012).

184. Tumanov, S. & Kamphorst, J. J. Recent advances in expanding the coverage of the lipidome. *Curr. Opin. Biotechnol.* **43,** 127–133 (2017).

185. Saha, S., Shalova, I. N. & Biswas, S. K. Metabolic regulation of macrophage phenotype and function. *Immunol. Rev.* **280,** 102–111 (2017).

186. Ray, A. & Dittel, B. N. Isolation of Mouse Peritoneal Cavity Cells. *J. Vis. Exp.* 9–11 (2010). doi:10.3791/1488

187. Misharin, A. V., Saber, R. & Perlman, H. Eosinophil contamination of thioglycollate-elicited peritoneal macrophage cultures skews the functional readouts of in vitro assays. *J. Leukoc. Biol.* **92,** 325–331 (2012).

188. Norris, P. C., Reichart, D., Dumlao, D. S., Glass, C. K. & Dennis, E. A. Specificity of eicosanoid production depends on the TLR-4-stimulated macrophage phenotype. *J. Leukoc. Biol.* **90,** 563–574 (2011).

189. Ekroos, K. & Ed, R. *Lipidomics*. (Wiley-VCH Verlag GmbH & Co. KGaA, 2012).

190. Schuhmann, K. *et al.* Shotgun lipidomics on a LTQ Orbitrap mass spectrometer by successive switching between acquisition polarity modes. *J. Mass Spectrom.* **47,** 96–104 (2012).

191. Lipidomics, S., Orbitrap, L. T. Q., Spectrometer, M., Switching, S. & Modes, A. P. for the manuscript by Schuhmann et al " Shotgun Lipidomics on a LTQ Orbitrap".

192. O'Connor, A. *et al.* LipidFinder: A computational workflow for discovery of lipids identifies eicosanoid-phosphoinositides in platelets. *JCI Insight* **2,** (2017).

193. Prince, J. T. & Marcotte, E. M. Chromatographic alignment of ESI-LC-MS proteomics data sets by ordered bijective interpolated warping. *Anal. Chem.* **78,** 6140–6152 (2006).

194. Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R. & Siuzdak, G. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **78,** 779–787 (2006).

195. Zamboni, N., Saghatelian, A. & Patti, G. J. Defining the Metabolome: Size, Flux, and Regulation. *Mol. Cell* **58,** 699–706 (2015).

196. Walker, J. M. *Lipidomics*. (Humana Press, 2010).

197. Perera, D. *et al.* Differential DNA repair underlies mutation hotspots at active promoters in cancer genomes. *Nature* **532,** 259–63 (2016).

198. Wammers, M. *et al.* Reprogramming of pro-inflammatory human macrophages to an anti-inflammatory phenotype by bile acids. *Sci. Rep.* **8,** 255 (2018).

199. Li, Y.-H. *et al.* Inhibition of Macrophage Proinflammatory Cytokine Expression by Steroids and Recombinant IL-10. *Neonatology* **80,** 124–132 (2001).

200. Vernon-Roberts, B. in *Journal of cardiovascular pharmacology* **6,** 131–159 (1969).

201. Roxas, B. A. & Li, Q. Significance analysis of microarray for relative quantitation of LC/MS data in proteomics. *BMC Bioinformatics* **9,** 187 (2008).

202. Chong, J. *et al.* MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **46,** W486–W494 (2018).

203. Ruxton, G. D. The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. *Behav. Ecol.* **17,** 688–690 (2006).

204. Rasmiena, A. A., Ng, T. W. & Meikle, P. J. Metabolomics and ischaemic heart disease. *Clin. Sci.* **124,** 289–306 (2013).

205. Wagner, S., Scholz, K., Sieber, M., Kellert, M. & Voelkel, W. Tools in metabonomics: An integrated validation approach for LC-MS metabolic profiling of mercapturic acids in human urine. *Anal. Chem.* **79,** 2918–2926 (2007).

206. Sumner, L. W. *et al.* Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **3,** 211–221 (2007).

207. Breslow, D. K. & Weissman, J. S. Membranes in balance: mechanisms of sphingolipid homeostasis. *Mol. Cell* **40,** 267–79 (2010).

208. Hannun, Y. A. & Obeid, L. M. Many ceramides. *J. Biol. Chem.* **286,** 27855–27862 (2011).

209. Mauerer, R., Ebert, S. & Langmann, T. High glucose, unsaturated and saturated fatty acids differentially regulate expression of ATP-binding cassette transporters ABCA1 and ABCG1 in human macrophages. *Exp. Mol. Med.* **41,** 126–132 (2009).

210. Nixon, G. F. Sphingolipids in inflammation: Pathological implications and potential therapeutic targets. *Br. J. Pharmacol.* **158,** 982–993 (2009).

211. Aderem, A. & Underhill, D. M. Mechanism of Phagocytosis in Macrophages. *Annu. Rev. Immunol.* **17,** 593–623 (1999).

212. Nakagawa, Y. The metabolism of glycerophospholipids and its regulation of macrophages. *Tanpakushitsu Kakusan Koso.* **36,** 222–34 (1991).

213. Poelma, D. L. H., Lachmann, B., Haitsma, J. J., Zimmermann, L. J. & van Iwaarden, J. F. Influence of phosphatidylglycerol on the uptake of liposomes by alveolar cells and on lung function. *J. Appl. Physiol.* **98,** 1784–1791 (2005).

214. Maceyka, M., Milstien, S. & Spiegel, S. Sphingosine-1-phosphate : the Swiss army knife of sphingolipid signaling. *J. Lipid Res.* (2009).

215. Monick, M. M. *et al.* Cooperative Prosurvival Activity by ERK and Akt in Human Alveolar Macrophages is Dependent on High Levels of Acid Ceramidase Activity. *J. Immunol.* **173,** 123–135 (2004).

216. Gutierrez, M. G., Gonzalez, A. P., Anes, E. & Griffiths, G. Role of lipids in killing mycobacteria by macrophages: Evidence for NF-κB-dependent and -independent killing induced by different lipids. *Cell. Microbiol.* **11,** 406–420 (2009).

217. Grabowski, G. A. Delivery of lysosomal enzymes for therapeutic use: glucocerebrosidase as an example. *Expert Opin. Drug Deliv.* **3,** 771–782 (2006).

218. Schaade, L., Ritter, K., Schiebel, H. M., Thomssen, R. & Kleines, M. Characterization of cytostatically active glycosphingolipids isolated from thioglycollate-elicited murine macrophages. *IUBMB Life* **48,** 353–358 (1999).

219. Stamatos, N. M. *et al.* Differential expression of endogenous sialidases of human monocytes during cellular differentiation into macrophages. *FEBS J.* **272,** 2545–2556 (2005).

220. Berenson, C. S., Gallery, M. A., Smigiera, J. M. & Rasp, R. H. The role of ceramide of human macrophage gangliosides in activation of human macrophages. *J. Leukoc. Biol.* **72,** 492–502 (2002).

221. Balsinde, J., Balboa, M. A. & Dennis, E. A. Inflammatory activation of arachidonic acid signaling in murine P388D1 macrophages via sphingomyelin synthesis. *J. Biol. Chem.* **272,** 20373–20377 (1997).

222. Chambard, J. C., Lefloch, R., Pouysségur, J. & Lenormand, P. ERK implication in cell cycle regulation. *Biochim. Biophys. Acta - Mol. Cell Res.* **1773,** 1299–1310 (2007).

223. Cagnol, S. & Chambard, J. C. ERK and cell death: Mechanisms of ERK-induced cell death - Apoptosis, autophagy and senescence. *FEBS J.* **277,** 2–21 (2010).

224. Im, J. I. K., Im, S. A. N. G. H. Y. U. N. K., Ohnson, V. I. J. J. & Harma, R. A. P. S. Extracellular Signal-Regulated Kinase-Signaling-Dependent G2 / M Arrest and Cell Death in Murine Macrophages By Cadmium. **24,** 3069–3077 (2005).

225. Filippi-Chiela, E. C., Villodre, E. S., Zamin, L. L. & Lenz, G. Autophagy interplay with apoptosis and cell cycle regulation in the growth inhibiting effect of resveratrol in glioma cells. *PLoS One* **6,** (2011).

226. Krank, J., Murphy, R. C., Barkley, R. M., Duchoslav, E. & McAnoy, A. LIPID MAPS mass spectrometry methods chapters. *Methods Enzymol.* **432,** 1–20 (2007).

227. Murphy, R. C. & Jr, A. H. M. Biochimica et Biophysica Acta Lipidomics and Imaging Mass Spectrometry. *BBA - Mol. Cell Biol. Lipids* **1811,** 635–636 (2011).

228. Scherer, M., Leuthäuser-Jaschinski, K., Ecker, J., Schmitz, G. & Liebisch, G. A rapid and quantitative LC-MS/MS method to profile sphingolipids. *J. Lipid Res.* **51,** 2001–2011 (2010).

229. Merrill, B. A. H., Caligan, T. B. & Wang, E. *[l] Analysis of Sphingoid Bases and Sphingoid Liquid Chromatography*. *Sphingolipid Metabolism and Cell Signaling* **312,** (Elsevier Masson SAS, 2000).

230. Mullen, T. D., Hannun, Y. A. & Obeid, L. M. Ceramide synthases at the centre of sphingolipid metabolism and biology. *Biochem. J.* **441,** 789–802 (2012).

231. Ternes, P., Brouwers, J. F. H. M., van den Dikkenberg, J. & Holthuis, J. C. M. Sphingomyelin synthase SMS2 displays dual activity as ceramide phosphoethanolamine synthase. *J. Lipid Res.* **50,** 2270–2277 (2009).

232. Andreyev, A. Y. *et al.* Subcellular organelle lipidomics in TLR-4-activated macrophages,. *J. Lipid Res.* **51,** 2785–2797 (2010).

233. Kolter, T. & Sandhoff, K. Lysosomal degradation of membrane lipids. *FEBS Lett.* **584,** 1700–1712 (2010).

234. Van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: Where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **9,** 112–124 (2008).

235. Sezgin, E., Levental, I., Mayor, S. & Eggeling, C. The mystery of membrane organization: Composition, regulation and roles of lipid rafts. *Nat. Rev. Mol. Cell Biol.* **18,** 361–374 (2017).

236. Harayama, T. & Riezman, H. Understanding the diversity of membrane lipid composition. *Nat. Rev. Mol. Cell Biol.* **19,** 281–296 (2018).

237. Tulone, C., Tsang, J., Prokopowicz, Z., Grosvenor, N. & Chain, B. Natural cathepsin E deficiency in the immune system of C57BL/6J mice. *Immunogenetics* **59,** 927–935 (2007).

238. Kolter, T. & Sandhoff, K. Recent advances in the biochemistry of sphingolipidoses. *Brain Pathol.* **8,** 79–100 (1998).

239. McCollister, B. D., Myers, J. T., Jones-Carson, J., Voelker, D. R. & Vazquéz-Torres, A. Constitutive acid sphingomyelinase enhances early and late macrophage killing of Salmonella enterica serovar typhimurium. *Infect. Immun.* **75,** 5346–5352 (2007).

240. Borger, D. K., Sidransky, E. & Aflaki, E. New macrophage models of Gaucher disease offer new tools for drug development. *Macrophage* **2,** e712 (2015).

241. Kolter, T. & Sandhoff, K. Sphingolipid metabolism diseases. *Biochim. Biophys. Acta - Biomembr.* **1758,** 2057–2079 (2006).

242. Jenkins, R. W., Canals, D. & Hannun, Y. A. Roles and regulation of secretory and lysosomal acid sphingomyelinase. *Cell. Signal.* **21,** 836–846 (2009).

243. Körschen, H. G. *et al.* The non-lysosomal β-glucosidase GBA2 is a non-integral membrane-associated protein at the endoplasmic reticulum (ER) and Golgi. *J. Biol. Chem.* **288,** 3381–3393 (2013).

244. Muto, N., Yamamoto, M., Tani, S. & Yonezawa, S. Characteristic distribution of cathepsin e which immunologically cross-reacts with the 86-kda acid proteinase from rat gastric mucosa. *J. Biochem.* **103,** 629–632 (1988).

245. Zaidi, N. & Kalbacher, H. in *Handbook of Proteolytic Enzymes* 42–49 (Elsevier, 2013). doi:10.1016/B978-0-12-382219-2.00006-5

246. Cook, M., Caswell, R. C., Richards, R. J., Kay, J. & Tatnell, P. J. Regulation of human and mouse procathepsin E gene expression. *Eur. J. Biochem.* **268,** 2658–68 (2001).

247. Chen, P. M., Gombart, Z. J. & Chen, J. W. Chloroquine treatment of ARPE-19 cells leads to lysosome dilation and intracellular lipid accumulation: possible implications of lysosomal dysfunction in macular degeneration. *Cell Biosci.* **1,** 10 (2011).

248. Abbas-Terki, T., Blanco-Bose, W., Déglon, N., Pralong, W. & Aebischer, P. Lentiviral-mediated RNA interference. *Hum. Gene Ther.* **13,** 2197–201 (2002).

249. Meister, G. & Tuschi, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* **431,** 343–349 (2004).

250. N. Taniguchi, A. Suzuki, Y. Ito, H. Narimatsu, T. Kawasaki, S. H. *Experimental Glycoscience*. (Springer, 2008).

251. Osmani, N., Pontabry, J., Comelles, J. & Fekonja, N. An Arf6- and caveolae-dependent pathway links hemidesmosome remodeling and mechanoresponse. 7–9 (2017).

252. Brady, R. 0, Kanfer, J. N., Mock, M. B. & Fredrickson, D. S. The metabolism of sphingomyelin. II. Evidence of an enzymatic deficiency in Niemann-Pick disease. *Proc Natl Acad Sci USA* **55,** 366–369 (1966).

253. Fukasawa, M., Nishijima, M., Itabe, H., Takano, T. & Hanada, K. Reduction of Sphingomyelin Level without Accumulation of Ceramide in Chinese Hamster Ovary Cells Affects Detergent-resistant Membrane Domains and Enhances Cellular Cholesterol Efflux to Methyl-β-cyclodextrin. *J. Biol. Chem.* **275,** 34028–34034 (2000).

254. Yan, N., Ding, T., Dong, J., Li, Y. & Wu, M. Sphingomyelin synthase overexpression increases cholesterol accumulation and decreases cholesterol secretion in liver cells. 1–8 (2011).

255. Asano S, Kitatani K, Taniguchi M, et al. Regulation of cell migration by sphingomyelin synthases: sphingomyelin in lipid rafts decreases responsiveness to signaling by the CXCL12/CXCR4 pathway. *Mol Cell Biol.* **32,** 3242–52 (2012).

256. Vida, C. *et al.* Role of macrophages in age-related oxidative stress and lipofuscin accumulation in mice. *Redox Biol.* **12,** 423–437 (2017).

257. Sorli, S. *et al.* The nonlysosomal ␣ -glucosidase GBA2 promotes endoplasmic reticulum stress and impairs tumorigenicity of human melanoma cells. 489–498.

258. Gonzalez-Carmona, M. A. *et al.* Beta-glucosidase 2 knockout mice with increased glucosylceramide show impaired liver regeneration. *Liver Int.* **32,** 1354–1362 (2012).

259. Yildiz, Y. *et al.* Mutation of β -glucosidase 2 causes glycolipid storage disease and impaired male fertility. **116,** (2006).

260. Pandey, M. K. *et al.* Complement drives glucosylceramide accumulation and tissue inflammation in Gaucher disease. *Nat. Publ. Gr.* **543,** 108–112 (2017).

261. Zaidi, N., Hermann, C., Herrmann, T. & Kalbacher, H. Biochemical and Biophysical Research Communications Emerging functional roles of cathepsin E. *Biochem. Biophys. Res. Commun.* **377,** 327–330 (2008).

262. Tsukuba, T. *et al.* Cathepsin E Deficiency Impairs Autophagic Proteolysis in Macrophages. *PLoS One* **8,** e82415 (2013).

263. Heath, V. C/EBP deficiency results in hyperproliferation of hematopoietic progenitor cells and disrupts macrophage development in vitro and in vivo. *Blood* **104,** 1639–1647 (2004).

264. Scott, C. L. *et al.* Bone marrow-derived monocytes give rise to self-renewing and fully differentiated Kupffer cells. *Nat. Commun.* **7,** 10321 (2016).

265. Morales, A., Lee, H., Goñi, F. M., Kolesnick, R. & Fernandez-Checa, J. C. Sphingolipids and cell death. *Apoptosis* **12,** 923–939 (2007).

266. Bektas, M. & Spiegel, S. Glycosphingolipids and cell death. *Glycoconj. J.* **20,** 39–47 (2004).

267. Reis, A. *et al.* A comparison of five lipid extraction solvent systems for lipidomic studies of human LDL. *J. Lipid Res.* **54,** 1812–1824 (2013).

268. Serbulea, V. *et al.* Macrophage phenotype and bioenergetics are controlled by oxidized phospholipids identified in lean and obese adipose tissue. *Proc. Natl. Acad. Sci. U. S. A.* **115,** E6254–E6263 (2018).

269. Namgaladze, D. & Brüne, B. Macrophage fatty acid oxidation and its roles in macrophage polarization and fatty acid-induced inflammation. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1861,** 1796–1807 (2016).

270. Mallat, Z. Macrophages. *Arterioscler. Thromb. Vasc. Biol.* **34,** 2509–19 (2014).

271. Liddiard, K., Rosas, M., Davies, L. C., Jones, S. a & Taylor, P. R. Macrophage heterogeneity and acute inflammation. *Eur. J. Immunol.* **41,** 2503–8 (2011).

272. Davies, L. C. *et al.* A quantifiable proliferative burst of tissue macrophages restores homeostatic macrophage populations after acute inflammation. *Eur. J. Immunol.* **41,** 2155–64 (2011).

273. Hashimoto, D. *et al.* Article Tissue-Resident Macrophages Self-Maintain Locally throughout Adult Life with Minimal Contribution from Circulating Monocytes. *Immunity* **38,** 792–804 (1968).

274. Murray, P. J. *et al.* Macrophage Activation and Polarization: Nomenclature and Experimental Guidelines. *Immunity* **41,** 14–20 (2014).

275. Davies, L. C. *et al.* Peritoneal tissue-resident macrophages are metabolically poised to engage microbes using tissue-niche fuels. *Nat. Commun.* **8,** 2074 (2017).

276. Romero, I. & Bast, R. C. Minireview: human ovarian cancer: biology, current management, and paths to personalizing therapy. *Endocrinology* **153,** 1593–602 (2012).

277. Ruffell, B., Affara, N. I. & Coussens, L. M. Differential macrophage programming in the tumor microenvironment. *Trends Immunol.* **33,** 119–26 (2012).

278. Robinson-Smith, T. M. *et al.* Macrophages mediate inflammation-enhanced metastasis of ovarian tumors in mice. *Cancer Res.* **67,** 5708–16 (2007).

279. Moughon, D. L. *et al.* Macrophage Blockade Using CSF1R Inhibitors Reverses the Vascular Leakage Underlying Malignant Ascites in Late-Stage Epithelial Ovarian Cancer. *Cancer Res.* **75,** 4742–52 (2015).

280. Belaguli, N. S. *et al.* GATA6 Promotes Colon Cancer Cell Invasion by Regulating Urokinase Plasminogen Activator Gene Expression. *Neoplasia* **12,** 856-IN1 (2010).

281. Duan, R. D. & Nilsson, Å. Metabolism of sphingolipids in the gut and its relation to inflammation and cancer development. *Prog. Lipid Res.* **48,** 62–72 (2009).

282. Giussani, P., Tringali, C., Riboni, L., Viani, P. & Venerando, B. Sphingolipids: Key regulators of apoptosis and pivotal players in cancer drug resistance. *Int. J. Mol. Sci.* **15,** 4356–4392 (2014).

283. Ogretmen, B. Sphingolipid metabolism in cancer signalling and therapy. *Nat. Rev. Cancer* (2017).